THESIS / THÈSE

MASTER EN SCIENCES MATHÉMATIQUES

Les algorithmes de l'extragradient étendus aux problèmes d'équilibre

Jérouville, Samuel

Award date:
2007

FUNDP
Faculté des Sciences
Département de Mathématique

Rempart de la Vierge, 8
B-5000 Namur Belgique

Les algorithmes de l'extragradient étendus aux problèmes d'équilibre

Mémoire présenté pour l'obtention
du grade de
Licencié en Sciences Mathématiques
par

Samuel Jérouville

Promoteur : Prof. J.-J. STRODIOT

Année Académique 2006-2007

Résumé : L'objectif de ce mémoire est d'utiliser la technique du problème auxiliaire pour développer des algorithmes itératifs résolvant des problèmes d'équilibre. Le premier de ces algorithmes consiste en une extension de celui de l'extragradient aux problèmes d'équilibre. Dans cet algorithme, on n'exigera pas que la bifonction d'équilibre satisfasse la propriété de monotonicité, mais elle devra satisfaire une certaine condition de type Lipschitz. Pour éviter cette exigence, nous proposerons des recherches linéaires communément utilisées dans les inéquations variationnelles pour obtenir des algorithmes de type projection résolvant des problèmes d'équilibre.

Abstract : In this work we use the auxiliary problem technique to develop iterative schemes for solving equilibrium problems. The first one extends the extragradient algorithm to equilibrium problems. In this scheme we will not require that the equilibrium bifunction satisfy the property of monotonicity, but it will have to satisfy a certain Lipschitz-type condition. In order to avoid this requirement, we will suggest linesearches commonly used for the variational inequalities such that we will get projection-type algorithms for solving equilibrium problems.

Table des matières

Introduction.

L'objectif de ce mémoire est de résoudre des problèmes d'équilibre au sens de Blum et Oettli [2]. Pour ce faire, on utilise le principe du problème auxiliaire introduit par Mastroeni dans [5].
Le premier algorithme est une extension aux problèmes d'équilibre de l'algorithme de l'extragradient tel que présenté dans [6].
Dans cet algorithme, on n'exige pas que la bifonction d'équilibre satisfasse la propriété de monotonicité, mais plutôt qu'elle satisfasse une certaine condition de type Lipschitz.
Pour éviter cette hypothèse assez contraignante, on suggère d'utiliser des recherches linéaires de type Armijo le long des directions obtenues. Cette technique est semblable à celle utilisée dans [4] pour résoudre des inéquations variationnelles.
On examine ensuite le cas particulier des inéquations variationnelles mixtes.
Enfin, on étudie un problème particulier d'équilibre.

Le mémoire s'organise commme suit : dans un premier chapitre, on rappelle les méthodes classiques du point fixe, de l'extragradient et de projection sur un hyperplan.
Le second chapitre est consacré aux problèmes d'équilibre. On y présente le principe du problème auxiliaire et son application aux inéquations variationnelles généralisées.
Dans le troisième et dernier chapitre, on étend la méthode de l'extragradient aux problèmes d'équilibre.

Ce mémoire est basé sur le livre [4] et sur les articles [5] et [3].

Chapitre 1

Les méthodes de projection.

Les méthodes de projection sont des méthodes conceptuellement simples pour la solution d'inéquations variationnelles VI(K,F)

$$\text{trouver } x^* \in K \text{ tel que } \langle F(x^*), x - x^* \rangle \geq 0 \ \forall x \in K.$$

Il y a deux caractéristiques communes à toutes les méthodes de ce type. La première étant que leur implémentation requiert la capacité de calculer efficacement la projection sur l'ensemble K fermé et convexe, ce qui limite l'applicabilité des méthodes, spécialement quand une telle projection est numériquement chère ou/et difficile ; la seconde étant que la méthode ne requiert pas l'utilisation des dérivées de F et n'inclut aucune calcul numérique complexe, hormis la projection sur K. D'une part, cette dernière caractéristique rend ces méthodes extrêmement simples et numériquement bon marché lorsque la projection est facilement implémentable ; d'autre part, la non utilisation de l'information contenue dans les dérivées empêche ces méthodes de converger rapidement.

Par conséquent, pour des ensembles K tels que la projection Π_K est facilement calculée, les méthodes de projection peuvent être appliquées pour obtenir la solution de très nombreux problèmes du fait de leur simplicité ; elles peuvent aussi être utilisées dans un schéma en deux phases dans lequel une méthode de projection peu coûteuse numériquement est d'abord appliquée afin d'atteindre une région "prometteuse" où l'on passe à une méthode plus chère, mais plus rapide et plus précise. Dans la prochaine section, nous introduisons et analysons différentes variantes de la méthode basique de projection.

Mais avant cela, rappelons quelques définitions bien connues en optimisation qui nous seront utiles par la suite.

Définition 1.0.1 *Soient K un sous-ensemble non vide, fermé et convexe de $\mathbb{R}^n$ et F un opérateur de K dans $\mathbb{R}^n$. Ce dernier est dit*

1. *fortement monotone de module $\mu > 0$ si et seulement si $\forall x, y \in K$, nous avons*
$$\langle F(x) - F(y), x - y \rangle \geq \mu \|x - y\|^2;$$
2. *strictement monotone si et seulement si $\forall x, y \in K$ distincts, nous avons*
$$\langle F(x) - F(y), x - y \rangle > 0;$$

3. *monotone si et seulement si* $\forall x, y \in K$, *nous avons*

$$\langle F(x) - F(y), x - y\rangle \geq 0;$$

4. *pseudo monotone si et seulement si* $\forall x, y \in K$, *nous avons*

$$\langle F(x), y - x\rangle \geq 0 \text{ implique } \langle F(y), x - y\rangle \leq 0;$$

5. *Lipschitz continu de module* L *si et seulement si* $\forall x, y \in K$, *nous avons*

$$\|F(x) - F(y)\| \leq L\|x - y\|.$$

1.1 L'itération basique du point fixe.

Nous commençons avec la méthode la plus basique de projection, qui découle d'une application du théorème du point fixe de Banach. Cette méthode simple a une portée limitée dans les applications pratiques, mais elle est le prototype de méthodes plus élaborées et elle donne une illustration claire de la perspective en point fixe qui est sous-jacente à toutes les méthodes de type projection que nous examinons dans cette section.

Dans la suite de cette section, K est un sous-ensemble non vide, fermé et convexe de $\mathbb{R}^n$ et F est un opérateur continu de K dans $\mathbb{R}^n$. On suppose en outre qu'on dispose d'une $n \times n$ matrice D qui est symétrique et définie positive.

Les solutions de l'VI(K,F) sont les solutions de l'équation naturelle

$$F^{nat}_{K,D}(x) \equiv x - \Pi_{K,D}[x - D^{-1}F(x)] = 0,$$

et vice versa, où $\Pi_{K,D}$ est le projecteur sur K défini par la matrice D; par définition, pour tout $x \in \mathbb{R}^n$, $\Pi_{K,D}(x)$ est la solution du problème de minimisation convexe en la variable y

$$\min_{y\in K} \frac{1}{2}\|y - x\|^2_D,$$

avec $\frac{1}{2}\|y - x\|^2_D := \frac{1}{2}\langle D(y - x), y - x\rangle$.

En effet, $x^* = \Pi_{K,D}[x^* - D^{-1}F(x^*)]$

$\Leftrightarrow x^*$ est solution optimale du problème

$\min_{y\in K} \frac{1}{2}\langle D[y - x^* + D^{-1}F(x^*)], y - x^* + D^{-1}F(x^*)\rangle$

$\Leftrightarrow 0 \in D[x^* - x^* + D^{-1}F(x^*)] + N_K(x^*)$,

avec $N_K(x^*) := \{w : \langle w, x - x^*\rangle \leq 0 \; \forall x \in K\}$,

en utilisant la condition nécessaire et suffisante d'optimalité en programmation convexe,

$\Leftrightarrow -F(x^*) \in N_K(x^*)$

$\Leftrightarrow \langle F(x^*), x - x^* \rangle \geq 0 \ \forall x \in K.$

On note au passage que, par [6], ce projecteur est non expansif pour la norme D, c'est-à-dire que $\|\Pi_{K,D}(x) - \Pi_{K,D}(y)\|_D \leq \|x - y\|_D$.

Par conséquent, les points fixes de l'opérateur

$$x \mapsto \Pi_{K,D}[x - D^{-1}F(x)]$$

sont les solutions de l'VI(K,F) et vice versa. D'où, si ce dernier opérateur est une contraction, nous pouvons utiliser l'algorithme du point fixe de Banach pour calculer une solution de l'VI(K,F) : c'est l'*algorithme basique de projection.*

Algorithme 1.1.1

Données : *$x^0 \in K$ et une $n \times n$ matrice D symétrique et définie positive.*

Pas 0 : *Poser $k := 0$.*

Pas 1 : *Si $x^k = \Pi_{K,D}[x^k - D^{-1}F(x^k)]$, stop.*

Pas 2 : *Poser $x^{k+1} := \Pi_{K,D}[x^k - D^{-1}F(x^k)]$ et $k := k + 1$; revenir au Pas 1.*

Le résultat qui suit donne des conditions suffisantes sur l'opérateur F et la matrice D pour assurer la convergence de cet algorithme.

Théorème 1.1.1 *Soit $F : K \to \mathbb{R}^n$, où K est un sous-ensemble fermé et convexe de $\mathbb{R}^n$. Supposons que L et μ sont telles que pour tous x et y dans K,*

$$\langle F(x) - F(y), x - y \rangle \geq \mu \|x - y\|_2^2$$

et

$$\|F(x) - F(y)\|_2 \leq L\|x - y\|_2.$$

Si

$$L^2 \lambda_{max}(D) < 2\mu \lambda_{min}^2(D),$$

l'opérateur $\Pi_{K,D}[x - D^{-1}F(x)]$ est une contraction de K dans K pour la norme D ; par conséquent, chaque suite $\{x^k\}$ produite par l'algorithme qui précède converge vers l'unique (puisque D est définie positive) solution de l'VI(K,F).

Preuve. D'après le théorème de Rayleigh, nous avons, pour toute matrice D symétrique et définie positive, que pour tous x et y dans $\mathbb{R}^n$,

$$\lambda_{min}(D)\|x - y\|_2^2 \leq \|x - y\|_D^2 \leq \lambda_{max}(D)\|x - y\|_2^2.$$

Donc, pour tous vecteurs x et y de K, nous pouvons écrire

$\|\Pi_{K,D}[x - D^{-1}F(x)] - \Pi_{K,D}[y - D^{-1}F(y)]\|_D^2$

$\leq \|(x-y) + D^{-1}[F(y) - F(x)]\|_D^2$ en utilisant la non expansivité de $\Pi_{K,D}$ pour la norme D

$= \|F(x) - F(y)\|_{D^{-1}}^2 + \|x-y\|_D^2 - 2\langle F(x) - F(y), x-y\rangle$ par définition de la norme D

$\leq [1 + \frac{L^2}{\lambda_{min}^2(D)} - \frac{2\mu}{\lambda_{max}(D)}]\|x-y\|_D^2$

puisque, d'une part,

$\|F(x) - F(y)\|_{D^{-1}}^2$

$\leq \frac{1}{\lambda_{min}(D)}\|F(x) - F(y)\|_2^2$ vu que D^{-1} est symétrique et définie positive et que $\lambda_{max}(D^{-1}) = \frac{1}{\lambda_{min}(D)}$

$\leq \frac{L^2}{\lambda_{min}(D)}\|x-y\|_2^2$ par hypothèse

$\leq \frac{L^2}{\lambda_{min}(D)}\frac{\|x-y\|_D^2}{\lambda_{min}(D)}$

$= \frac{L^2}{\lambda_{min}^2(D)}\|x-y\|_D^2$

et, d'autre part,

$-2\langle F(x) - F(y), x-y\rangle$

$\leq -2\mu\|x-y\|_2^2$ par hypothèse

$\leq \frac{-2\mu}{\lambda_{max}(D)}\|x-y\|_D^2.$

Par conséquent, en utilisant la dernière hypothèse, $\Pi_{K,D}[x - D^{-1}F(x)]$ est une contraction.

Si D est une matrice multiple de l'identité de coefficient τ strictement positif, la dernière hypothèse devient

$$\tau > \frac{L^2}{2\mu}.$$

En conséquence, pour une VI(K,F) donnée avec un opérateur F fortement monotone et lipschitz continu, pour autant que les constantes L et μ soient disponibles, et que τ est choisi de sorte à vérifier la condition décrite plus haut, la suite $\{x^k\}$ définie itérativement par

$$x^{k+1} \equiv \Pi_K[x^k - \tau^{-1}F(x^k)],\ k = 0, 1, 2, \ldots$$

converge vers l'unique solution de l'VI(K,F), pour tout x^0 de K.

Si F est le gradient de la fonction convexe θ avec un ensemble non vide de minima sous contraintes dans K, il est connu que la méthode décrite plus haut converge pour des valeurs de τ suffisamment grandes. Ceci peut nous amener à supposer qu'il serait possible de prouver la convergence de l'algorithme basique de projection sous une condition de monotonicité affaiblie sur F. Le prochain exemple montre que cette conjecture est fausse.

Exemple 1.1.1 *Soit* $F(x) := Ax$, *où*

$$A := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Notons que la matrice jacobienne de F est égale à la matrice asymétrique A, de sorte que F n'est pas un opérateur gradient. On voit immédiatement que l'unique solution de l'VI(F,$\mathbb{R}^2$)

$$\text{trouver } x^* \in K \text{ tel que } \langle F(x^*), y - x^* \rangle \geq 0 \ \forall y \in K$$

$$\Leftrightarrow \text{trouver } x^* \in K \text{ tel que } x_2^* y_1 - x_1^* y_2 \geq 0 \ \forall y \in K$$

est $x^* = 0$.

Cependant, un calcul direct montre que, pour tous $x^0 \neq 0$ et $\tau > 0$, nous avons, en utilisant la décomposition de la norme en somme de produits scalaires,

$$\|x^{k+1} - x^*\| = \|x^{k+1}\| = \|x^k - \tau^{-1} F(x^k)\| = \sqrt{1 + \tau^{-2}} \|x^k\| > \|x^k\| = \|x^k - x^*\|,$$

si bien que l'algorithme basique de projection ne peut clairement pas converger vers x^*.

L'un de nos objectifs dans le reste de cette section est de développer des algorithmes de projection qui, au contraire de l'algorithme basique de projection, ne requièrent ni la monotonicité forte de F, ni la connaissance a priori des constantes L et μ typiquement inconnues en pratique.

Dans l'algorithme qui va suivre, les deux constantes L et μ seront remplacées par une constante de co-coervité. Dans la prochaine sous-section, nous présenterons un algorithme de type projection applicable à une VI(K,F) pseudo monotone, qui requiert uniquement la connaissance de la constante de lipschitz L de la fonction F. Et dans la sous-section qui suivra, nous présenterons un algorithme de type projection qui utilise une recherche linéaire simple et qui est applicable à l'VI(K,F), et qui ne requiert pas que F soit Lipschitz continu.

Dans le reste de cette sous-section, nous présenterons une analyse raffinée de l'algorithme basique de projection qui n'est pas basée sur l'argument en point fixe, mais qui mène à un résultat de convergence amélioré. On appliquera alors cette analyse à un schéma de projection à pas variable dans lequel on permet à la longueur du pas τ de varier d'une itération à l'autre.

C'est une extension du dernier pas de l'algorithme basique de projection qui, lui, est à pas constant. L'algorithme à pas variable qui en résultera n'applique cependant pas une méthode de recherche linéaire, puisque la longueur du pas τ_k n'est pas déterminée par une routine de ce type. Il est à noter que dans le second pas de l'algorithme qui suit, algorithme auquel on se référera comme étant l'*algorithme de projection à pas variable*, nous utiliserons τ_k à la place de τ_k^{-1}, mais il s'agit uniquement là d'un choix de notation.

Algorithme 1.1.2

Données : $x^0 \in K$.

Pas 0 : *Poser* $k := 0$.

Pas 1 : *Si* $x^k = \Pi_K[x^k - \tau_k F(x^k)]$, *stop.*

Pas 2 : *Choisir* $\tau_k > 0$. *Poser* $x^{k+1} := \Pi_{K,D}[x^k - \tau_k F(x^k)]$ *et* $k := k+1$ *; revenir au Pas 1.*

Le choix de la suite de scalaires $\{\tau_k\}$ est crucial dans la réussite de cet algorithme. Dans ce qui va suivre, nous montrerons que si F est co-coercif de module $c > 0$, c'est-à-dire que si

$$\langle F(x) - F(y), x - y\rangle \geq c\|F(x) - F(y)\|_2^2 \; \forall x, y \in K,$$

alors $\{\tau_k\}$ peut être choisie de sorte que la suite $\{x^k\}$ converge vers une solution de l'VI(K,F). Puisque chaque opérateur fortement monotone et Lipschitz continu est co-coercif (en effet, si l'opérateur F est, d'une part, fortement monotone et, d'autre part, Lipschitz continu, $\forall x, y : \langle F(y) - F(x), y - x\rangle \geq \mu\|y - x\|^2 \geq \frac{\mu}{L^2}\|F(y) - F(x)\|^2$), mais non vice versa (un opérateur co-coercif de module c étant, après application du théorème de Cauchy-Schwarz, Lipschitz continu de module $\frac{1}{c}$; mais un opérateur constant, trivialement co-coercif, n'étant pas fortement monotone), nous aurons par conséquent étendu avec succès la convergence de l'algorithme de projection à une classe plus large d'VI(K,F).

Pour arriver à nos fins, nous avons besoin de deux résultats préliminaires, le premier ayant un intérêt indépendant. Spécifiquement, nous considérons l'application de l'algorithme qui précède au cas où $K = R^n$, c'est-à-dire à la solution du système d'équations $x = x - \tau_k F(x)$ pour un certain $\tau_k > 0$ $\Leftrightarrow F(x) = 0$. Dans ce cas, l'algorithme se réduit à l'itération

$$x^{k+1} := x^k - \tau_k F(x^k), \;\; k = 0, 1, 2, \ldots$$

En fait, on considère une situation plus générale où le terme de correction $\tau_k F(x^k)$ est remplacé par $F^k(x^k)$, où toutes les F^k ont le même ensemble de zéros. Cette itération généralisée prend alors la forme

$$x^{k+1} = x^k - F^k(x^k), \;\; k = 0, 1, 2, \ldots$$

Le second résultat établit la convergence de la suite $\{x^k\}$ générée.

Lemme 1.1.1 *Soient $F^k : I\!R^n \to I\!R^n$ pour $k = 0, 1, 2, \ldots$ des opérateurs co-coercifs sur $I\!R^n$ de modules c_k tels que*

$$\rho := inf_k \; c_k > \frac{1}{2}.$$

Si toutes les fonctions F^k ont le même ensemble non vide S de zéros et si

$$inf_k \; \|F^k(x)\| > 0 \;\; \forall x \notin S,$$

alors la suite $\{x^k\}$ produite par l'itération généralisée présentée plus haut converge vers un point x^* dans S.

Preuve. Soient $\{x^k\}$ et $\{y^k\}$ deux séquences produites par l'itération généralisée en partant de x^0 et y^0 respectivement. Nous montrons tout d'abord que

$$\lim_{k\to\infty} \|x^k - y^k\| =: \sigma < \infty;$$
$$\sum_{k=0}^{\infty} \|F^k(x^k) - F^k(y^k)\|^2 < \infty.$$

Prenant en compte la condition de co-coercivité, un calcul direct donne

$$0$$

$$\leq \|x^{k+1} - y^{k+1}\|^2$$

$$= \|x^k - F^k(x^k) - y^k + F^k(y^k)\|^2 \text{ par définition de } x^{k+1} \text{ et } y^{k+1}$$

$$= \|x^k - y^k\|^2 + \|F^k(x^k) - F^k(y^k)\|^2 - 2\langle F^k(x^k) - F^k(y^k), x^k - y^k\rangle$$

$$\leq \|x^k - y^k\|^2 - (2c_k - 1)\|F^k(x^k) - F^k(y^k)\|^2$$

$$\leq \|x^k - y^k\|^2 - (2\rho - 1)\|F^k(x^k) - F^k(y^k)\|^2 \text{ puisque } \rho := \inf_k c_k.$$

Ceci montre que la suite positive $\{\|x^k - y^k\|\}$ est décroissante (puisque $\rho > \frac{1}{2}$), et a donc une limite que nous appellerons σ ; ceci mène à la première inégalité annoncée. Pour montrer la seconde, il suffit d'observer qu'en appliquant récursivement l'inégalité calculée plus haut, on a

$$\|x^{k+1} - y^{k+1}\|^2 \leq \|x^0 - y^0\|^2 - (2\rho - 1)\textstyle\sum_{i=0}^{k} \|F^i(x^i) - F^i(y^i)\|^2$$

$$\Leftrightarrow \textstyle\sum_{i=0}^{k} \|F^i(x^i) - F^i(y^i)\|^2 \leq \frac{1}{2\rho-1}(\|x^0 - y^0\|^2 - \|x^{k+1} - y^{k+1}\|^2).$$

Soit $\bar{x}$ un élément arbitraire dans S. Si nous commençons le schéma itératif en $y^0 = \bar{x}$, nous avons $y^k = \bar{x}$ pour tout k (vu que $y^1 = y^0 - F^0(y^0) = \bar{x} - F^0(\bar{x}) = \bar{x}$ puisque $\bar{x} \in S$, et ainsi de suite). Par la seconde inégalité calculée, nous avons

$$\sum_{k=0}^{\infty} \|F^k(x^k) - F^k(\bar{x})\|^2 = \sum_{k=0}^{\infty} \|F^k(x^k)\|^2 < \infty,$$

alors que la première inégalité calculée prouve que

$$\lim_{k\to\infty} \|x^k - \bar{x}\|^2 =: \sigma$$

existe, de sorte que la suite $\{x^k\}$ est bornée. Soit x^* la limite de la sous-suite $\{x^k : k \in \kappa \subseteq \mathbb{N}\}$, qui est bien définie par le théorème de Bolzano-Weierstrass. Montrons que $x^* \in S$. Puisque F^k est co-coercif de module c_k, et donc Lipschitz continu de module $\frac{1}{c_k}$,

$$0 \leq \|F^k(x^k) - F^k(x^*)\| \leq \frac{1}{c_k}\|x^k - x^*\| \leq \frac{1}{\rho}\|x^k - x^*\| \text{ par définition de } \rho,$$

ce qui implique, en appliquant le théorème de l'étau,

$$\lim_{k(\in\kappa)\to\infty} \|F^k(x^k) - F^k(x^*)\| = 0.$$

Etant donné que $\sum_{k=0}^{\infty} \|F^k(x^k)\|^2 < \infty$, nous avons aussi

$$\lim_{k\to\infty} \|F^k(x^k)\| = 0,$$

ce qui implique, par l'unicité de la limite, que

$$\lim_{k(\in\kappa)\to\infty} \|F^k(x^k)\| = 0.$$

Ainsi, puisqu'en utilisant un artifice de calcul et l'inégalité triangulaire, $0 \leq \|F^k(x^*)\| \leq \|F^k(x^*) - F^k(x^k)\| + \|F^k(x^k)\|$,

$$\lim_{k(\in\kappa)\to\infty} \|F^k(x^*)\| = 0.$$

Grâce à cette limite et à la dernière hypothèse, nous pouvons conclure que $x^* \in S$. Utilisant la relation $\lim_{k\to\infty} \|x^k - \bar{x}\|^2 =: \sigma$ avec $\bar{x} = x^* \in S$, on peut affirmer que la suite $\{x^k\}$ est convergente ; puisqu'en outre, la sous-suite $\{x^k : k \in \kappa \subseteq \mathbb{N}\}$ converge vers x^*, l'unicité de la limite implique que la suite $\{x^k\}$ elle-même converge vers x^*.

Remarque 1.1.1 *L'hypothèse selon laquelle chaque F^k est co-coercif sur $\mathbb{R}^n$ a été émise afin de simplifier quelque peu la preuve qui précède. On peut aisément voir qu'il est suffisant de supposer que chaque F^k soit co-coercif sur un ensemble fermé contenant la suite $\{x^k\}$ et S.*

Supposons que $F^{nat}_{K,\tau}(\cdot) \equiv \cdot - \Pi_K[\cdot - \tau F(\cdot)]$ désigne l'opérateur naturel de l'VI(K,τ,F). Le lemme qui va suivre montre que $F^{nat}_{K,\tau}(\cdot)$ est un opérateur co-coercif sur K ; ce résultat est essentiel pour appliquer le dernier lemme dans l'analyse du dernier algorithme.

Lemme 1.1.2 *Supposons que le mapping $F : K \subseteq \mathbb{R}^n \to \mathbb{R}^n$, avec K convexe et non vide, soit co-coercif sur K de module c. Si $\tau \in]0, 4c[$, alors $F^{nat}_{K,\tau}(\cdot)$ est co-coercif sur K de constante $1 - \frac{\tau}{4c}$.*

Preuve. Puisque, par [6], la projection est co-coercive de constante 1, il s'ensuit que, pour tous vecteurs x et y,

$$\langle \Pi_K[x - \tau F(x)] - \Pi_K[y - \tau F(y)], [x - \tau F(x)] - [y - \tau F(y)] \rangle$$

$\geq \|\Pi_K[x - \tau F(x)] - \Pi_K[y - \tau F(y)]\|^2.$

Après manipulation, on déduit que $\langle F^{nat}_{K,\tau}(x) - F^{nat}_{K,\tau}(y), x - y\rangle$

$\geq \|F^{nat}_{K,\tau}(x) - F^{nat}_{K,\tau}(y)\|^2 + \tau\langle F(x) - F(y), x - y\rangle - \tau\langle F^{nat}_{K,\tau}(x) - F^{nat}_{K,\tau}(y), F(x) - F(y)\rangle$

étant donné que

1. $\langle F^{nat}_{K,\tau}(x) - F^{nat}_{K,\tau}(y), x - y\rangle$

   $= \langle x - y - \Pi_K[x - \tau F(x)] + \Pi_K[y - \tau F(y)], x - y\rangle$

   $= \|x - y\|^2 - \langle \Pi_K[x - \tau F(x)] - \Pi_K[y - \tau F(y)], x - y\rangle$
2. $\|F^{nat}_{K,\tau}(x) - F^{nat}_{K,\tau}(y)\|^2$

   $= \|x - y - \Pi_K[x - \tau F(x)] + \Pi_K[y - \tau F(y)]\|^2$

   $= \|x - y\|^2 - 2\langle x - y, \Pi_K[x - \tau F(x)] - \Pi_K[y - \tau F(y)]\rangle + \|\Pi_K[x - \tau F(x)] - \Pi_K[y - \tau F(y)]\|^2$
3. $\tau\langle F(x) - F(y), x - y\rangle - \tau\langle F^{nat}_{K,\tau}(x) - F^{nat}_{K,\tau}(y), F(x) - F(y)\rangle$

   $= \langle \tau F(x) - \tau F(y), x - y - x + \Pi_K[x - \tau F(x)] + y - \Pi_K[y - \tau F(y)]\rangle$

   $= \langle \tau F(x) - \tau F(y), \Pi_K[x - \tau F(x)] - \Pi_K[y - \tau F(y)]\rangle.$

Utilisant la co-coercivité de F, l'inégalité implique

$\langle F^{nat}_{K,\tau}(x) - F^{nat}_{K,\tau}(y), x - y\rangle$

$\geq \|F^{nat}_{K,\tau}(x) - F^{nat}_{K,\tau}(y)\|^2 + \tau c\|F(x) - F(y)\|^2 - \tau\langle F^{nat}_{K,\tau}(x) - F^{nat}_{K,\tau}(y), F(x) - F(y)\rangle$

$= (1 - \frac{\tau}{4c})\|[F^{nat}_{K,\tau}(x) - F^{nat}_{K,\tau}(y)]\|^2$ puisque $\|\sqrt{\frac{\tau}{4c}}[F^{nat}_{K,\tau}(x) - F^{nat}_{K,\tau}(y)] - \sqrt{\tau c}[F(x) - F(y)]\|^2 \geq 0,$

ce qui établit le lemme.

On peut alors montrer qu'en combinant les deux derniers lemmes, on obtient la convergence de l'algorithme de projection à pas variable pour la classe des VI(K,F) co-coercives.

Théorème 1.1.2 *Soient K un sous-ensemble fermé et convexe de $\mathbb{R}^n$ et F un opérateur de K dans $\mathbb{R}^n$ qui est co-coercif sur K de module c. Supposons que SOL(K,F)$\neq \emptyset$. Si*

$$0 < inf_k\ \tau_k \leq sup_k\ \tau_k < 2c,$$

alors l'algorithme de projection à pas variable produit une suite $\{x^k\}$ convergeant vers une solution de l'VI(K,F).

Preuve. L'algorithme de projection à pas variable peut être vu comme étant une application de l'itération généralisée avec $F^k(x) := F^{nat}_{K,\tau_k}(x)$, dont l'ensemble des zéros, c'est-à-dire $\{x^* \mid x^* - \Pi_K[x^* - \tau_k F(x^*)] = 0\}$, coïncide clairement, comme déjà vu, avec SOL(K,F), c'est-à-dire $\{x^* \mid x^* = \Pi_K[x^* - \tau_k F(x^*)]\}$.

Par le dernier lemme, chaque F^k est co-coercif de module $1 - \frac{\tau_k}{4c}$ sur K qui est un ensemble fermé contenant la suite $\{x^k\}$ et SOL(K,F).

En outre, un simple argument montre que, pour tout x qui n'est pas solution de l'VI(K,F),

$$\inf_k \|F^k(x)\| > 0.$$

En passant à la contraposée, on doit prouver que

$$\text{si } \inf_k \|F^k(x)\| = 0, \text{ alors } x \in \text{ SOL}(K\ F).$$

Or, par définition de l'infimum, $\inf_k \|F^k(x)\| = 0 \Leftrightarrow \exists$ une sous-suite $\{F^{q_k}(x)\}$ telle que $\lim_{q_k \to \infty} \|F^{q_k}(x)\| = 0$.

La suite de pas correspondante $\{\tau_{q_k}\}$ vérifie quant à elle $\{\tau_{q_k}\} \subseteq [\inf_{\tau_k}, \sup_{\tau_k}]$ si bien que, par passage au théorème de Bolzano-Weierstrass et donc à une sous suite de $\{\tau_{q_k}\}$ et à son point d'adhérence τ, $\lim_{q_k \to \infty} \tau_{q_k} = \tau$ où la nouvelle sous-suite fille est, sans perte de généralité, indexée comme la sous-suite mère.

Ainsi, pour $\varepsilon > 0$ arbitraire, $\exists k_\varepsilon$ tel que $\forall k \geq k_\varepsilon$, on a $\|F^{q_k}(x)\| < \frac{\varepsilon}{2}$ et $|\tau_{q_k} - \tau| < \frac{\varepsilon}{2\|F(x)\|}$ en supposant pour l'instant que $\|F(x)\| \neq 0$.

Et donc,

$\|F^{nat}_{K,\tau}(x)\|$

$\leq \|F^{nat}_{K,\tau}(x) - F^{q_k}(x)\| + \|F^{q_k}(x)\|$

$= \|x - \Pi_K[x - \tau F(x)] - x + \Pi_K[x - \tau_{q_k} F(x)]\| + \|F^{q_k}(x)\|$

$\leq \| - x + \tau F(x) + x - \tau_{q_k} F(x)\| + \|F^{q_k}(x)\|$ puisque, par [6], le projecteur $\Pi_K(\cdot)$ est non expansif pour la norme deux

$\leq |\tau - \tau_{q_k}| \|F(x)\| + \|F^{q_k}(x)\|$

$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2}$

$= \varepsilon.$

Dans l'éventualité où $\|F(x)\| = 0$, le premier terme de la pénultième inégalité est nul, de sorte que l'inégalité finale tient toujours. Celle-ci implique que $\|F^{nat}_{K,\tau}(x)\| = 0$, c'est-à-dire que $F^{nat}_{K,\tau}(x) = 0$, et finalement que $x \in$ SOL(K,F).

Par conséquent, d'après la remarque vue plus tôt, on peut appliquer le lemme qui précède cette dernière pour déduire que la suite $\{x^k\}$ converge vers une solution de l'VI(K,F).

1.2 La méthode de l'extragradient.

Le dernier théorème est certainement une amélioration de celui qui le précède en ce sens où la monotonicité forte de F est remplacée par sa cocoercivité. Dans cette sous-section, nous introduisons un algorithme de projection qui exécute deux projections par itération. Bien que celui-ci requiert indubitablement le double de calculs numériques, le gain est significatif, car l'algorithme résultant est applicable à la classe des VI(K,F) pseudo monotones. Cependant, on requiert encore que l'opérateur F soit Lipschitz continu, et une estimation de sa constante de Lipschitz est nécessaire. La méthode de l'extragradient qu'on présente par après tient son nom de l'évaluation supplémentaire de F (et de la projection supplémentaire) que l'on effectue à chaque itération. Le nom tire son origine du cas d'une VI(K,F) symétrique. Dans ce cas, l'VI(k,F) représente la condition d'optimalité d'un problème d'optimisation différentiable, et l'évaluation supplémentaire de F correspond à une évaluation supplémentaire du gradient de la fonction objectif , d'où cette notion d'extragradient. Venons-en maintenant à l'algorithme de l'extragradient proprement dit.

Algorithme 1.2.1

Données : $x^0 \in K$ *et* $\tau > 0$.

Pas 0 : *Poser* $k := 0$.

Pas 1 : *Si* $x^k = \Pi_K[x^k - \tau F(x^k)]$, *stop.*

Pas 2 : *Calculer*

$$x^{k+\frac{1}{2}} := \Pi_K[x^k - \tau F(x^k)],$$

$$x^{k+1} := \Pi_K[x^k - \tau F(x^{k+\frac{1}{2}})];$$

et poser $k := k+1$ *; revenir au Pas 1.*

Afin de donner une motivation géométrique à cet algorithme, nous considérons son application à un système d'équations monotones (avec $K = \mathbb{R}^n$) ; dans ce cas, le Pas 2 devient

$$x^{k+\frac{1}{2}} := x^k - \tau F(x^k),$$

$$x^{k+1} := x^k - \tau F(x^{k+\frac{1}{2}}).$$

Considérons l'hyperplan $H^k := \{x \in \mathbb{R}^n : \langle F(x^{k+\frac{1}{2}}), x - x^{k+\frac{1}{2}}\rangle = 0\}$ et supposons que x^* soit une solution du système d'équations $F(x) = 0$, c'est-à-dire que $x^* \in \mathrm{SOL}(\mathbb{R}^n,F)$.
Par la monotonicité de F, nous avons $\langle F(x^{k+\frac{1}{2}}) - F(x^*), x^{k+\frac{1}{2}} - x^*\rangle \geq 0 \Leftrightarrow$

$\langle F(x^{k+\frac{1}{2}}), x^* - x^{k+\frac{1}{2}}\rangle \leq 0$.
De plus, puisque $x^{k+\frac{1}{2}} = x^k - \tau F(x^k) \Leftrightarrow x^k - x^{k+\frac{1}{2}} = \tau F(x^k)$, d'une part, $\langle F(x^{k+\frac{1}{2}}), x^k - x^{k+\frac{1}{2}}\rangle = \tau\langle F(x^{k+\frac{1}{2}}), F(x^k)\rangle$ et, d'autre part, $x^{k+\frac{1}{2}}$ converge vers x^k quand τ tend vers zéro de sorte que, pour τ suffisamment petit, nous avons, par continuité de F, $\langle F(x^{k+\frac{1}{2}}), F(x^k)\rangle > 0$ (on ne peut avoir $F(x^k) = 0$, car sinon nous nous serions arrêté au Pas 1).
Ainsi, pour τ suffisamment petit, $\langle F(x^{k+\frac{1}{2}}), x^k - x^{k+\frac{1}{2}}\rangle = \tau\langle F(x^{k+\frac{1}{2}}), F(x^k)\rangle > 0$ et H^k sépare x^k de toute solution x^*.
Par conséquent, nous voyons que $x^{k+1} = x^k - \tau F(x^{k+\frac{1}{2}})$ est un déplacement depuis x^k dans la direction de la projection de ce dernier point sur l'hyperplan séparateur H^k si bien que, pour τ suffisamment petit, x^{k+1} est plus proche que x^k de n'importe quelle solution du système d'équations.
Cette propriété (qu'on nomme monotonicité de Féjer de $\{x^k\}$ par rapport à l'ensemble-solution) est la base de l'analyse de convergence.
Si $K \neq \mathbb{R}^n$, la monotonicité de Féjer de $\{x^k\}$ par rapport à l'ensemble-solution peut encore être montrée, grâce aux propriétés du projecteur.

Dans la discussion qui précède, la monotonicité de F a seulement été utilisée pour montrer que $\langle F(x^{k+\frac{1}{2}}), x^* - x^{k+\frac{1}{2}}\rangle \leq 0$ pour toute solution x^*. Ceci motive la définition qui suit.

Définition 1.2.1 *On dit que l'opérateur $F : K \to \mathbb{R}^n$ est pseudo monotone sur K par rapport à SOL(K,F) si et seulement si ce dernier ensemble est non vide et pour tout $x^* \in$ SOL(K,F), on a*

$$\langle F(x), x - x^*\rangle \geq 0 \ \forall x \in K.$$

Ce concept est un affaiblissement de celui de pseudo monotonicité de F sur K, puisque la relation $\forall x, y \in K : \langle F(x), y - x\rangle \geq 0 \Rightarrow \langle F(y), x - y\rangle \leq 0$ se particularise à $\forall y \in K : \langle F(x^*), y - x^*\rangle \geq 0 \Rightarrow \langle F(y), y - x^*\rangle \geq 0$, et joue un rôle central dans la convergence de l'algorithme de l'extragradient. Le lemme-clé qui suit rend plus précise la précédente discussion sur l'algorithme et l'étend au cas $K \neq \mathbb{R}^n$.

Lemme 1.2.1 *Soient K un ensemble fermé et convexe dans $\mathbb{R}^n$ et F un opérateur de K dans $\mathbb{R}^n$ qui est pseudo monotone sur K par rapport à SOL(K,F) et Lipschitz continu sur K de module $L > 0$. Si x^* est une solution quelconque de l'VI(K,F), alors, pour tout k, on a*

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2 - (1 - \tau^2 L^2)\|x^{k+\frac{1}{2}} - x^k\|^2.$$

Preuve. Puisque $x^* \in$ SOL(K,F), $x^{k+\frac{1}{2}} \in K$, et F est pseudo monotone par rapport à SOL(K,F), nous avons $\langle F(x^{k+\frac{1}{2}}), x^{k+\frac{1}{2}} - x^*\rangle \geq 0$ pour tout k. Ceci implique, en utilisant un artifice de calcul, que

$$\langle F(x^{k+\frac{1}{2}}), x^* - x^{k+1}\rangle \leq \langle F(x^{k+\frac{1}{2}}), x^{k+\frac{1}{2}} - x^{k+1}\rangle.$$

En outre, on a

$\langle x^{k+1} - x^{k+\frac{1}{2}}, x^k - \tau F(x^{k+\frac{1}{2}}) - x^{k+\frac{1}{2}}\rangle$

$= \langle x^{k+1} - x^{k+\frac{1}{2}}, x^k - \tau F(x^k) - x^{k+\frac{1}{2}}\rangle + \tau\langle x^{k+1} - x^{k+\frac{1}{2}}, F(x^k) - F(x^{k+\frac{1}{2}})\rangle$ en utilisant un artifice de calcul

$= \langle x^{k+1} - \Pi_K[x^k - \tau F(x^k)], x^k - \tau F(x^k) - x^{k+\frac{1}{2}}\rangle + \tau\langle x^{k+1} - \Pi_K[x^k - \tau F(x^k)], F(x^k) - F(x^{k+\frac{1}{2}})\rangle$ par définition de $x^{k+\frac{1}{2}}$

$\leq \tau\langle x^{k+1} - \Pi_K[x^k - \tau F(x^k)], F(x^k) - F(x^{k+\frac{1}{2}})\rangle$ par la caractérisation géométrique de la projection orthogonale.

En posant $y^k := x^k - \tau F(x^{k+\frac{1}{2}})$ pour alléger les notations et en utilisant les inégalités qui précèdent, on obtient

$\|x^{k+1} - x^*\|^2$

$= \|\Pi_K(y^k) - y^k + y^k - x^*\|^2$ par définition de x^{k+1} et passage à un artifice de calcul

$= \|\Pi_K(y^k) - y^k\|^2 + \|y^k - x^*\|^2 + 2\langle\Pi_K(y^k) - y^k, y^k - x^*\rangle$

$\leq -\|\Pi_K(y^k) - y^k\|^2 + \|y^k - x^*\|^2$ par la caractérisation géométrique de la projection orthogonale.

$= -\|x^k - x^{k+1} - \tau F(x^{k+\frac{1}{2}})\|^2 + \|x^k - x^* - \tau F(x^{k+\frac{1}{2}})\|^2$ par les définitions de x^{k+1} et y^k

$= -\|x^k - x^{k+1}\|^2 - \|\tau F(x^{k+\frac{1}{2}})\|^2 + 2\tau\langle x^k - x^{k+1}, F(x^{k+\frac{1}{2}})\rangle + \|x^k - x^*\|^2 + \|\tau F(x^{k+\frac{1}{2}})\|^2 - 2\tau\langle x^k - x^*, F(x^{k+\frac{1}{2}})\rangle$

$= \|x^k - x^*\|^2 - \|x^k - x^{k+1}\|^2 + 2\tau\langle x^* - x^{k+1}, F(x^{k+\frac{1}{2}})\rangle$

$\leq \|x^k - x^*\|^2 - \|x^k - x^{k+1}\|^2 + 2\tau\langle x^{k+\frac{1}{2}} - x^{k+1}, F(x^{k+\frac{1}{2}})\rangle$ par la première inégalité démontrée plus haut

$= \|x^k - x^*\|^2 - \|x^k - x^{k+\frac{1}{2}}\|^2 - \|x^{k+\frac{1}{2}} - x^{k+1}\|^2 + 2\langle x^{k+1} - x^{k+\frac{1}{2}}, x^k - \tau F(x^{k+\frac{1}{2}}) - x^{k+\frac{1}{2}}\rangle$ en utilisant un artifice de calcul puis en réarrangeant quelque peu les termes

$\leq \|x^k - x^*\|^2 - \|x^k - x^{k+\frac{1}{2}}\|^2 - \|x^{k+\frac{1}{2}} - x^{k+1}\|^2 + 2\tau L\|x^{k+1} - x^{k+\frac{1}{2}}\|\|x^k - x^{k+\frac{1}{2}}\|$

par la seconde inégalité démontrée plus haut, par l'inégalité de Cauchy-Schwarz et par la Lipschitz continuité de F.

$\leq \|x^k - x^*\|^2 - (1 - \tau^2 L^2)\|x^k - x^{k+\frac{1}{2}}\|^2$ puisque $(\tau L\|x^k - x^{k+\frac{1}{2}}\| - \|x^{k+1} - x^{k+\frac{1}{2}}\|)^2 \geq 0$, ce qui achève la preuve du lemme.

Armé de ce lemme, nous pouvons établir la convergence de l'algorithme de l'extragradient. Notons que la constante de Lipschitz L de F joue toujours un rôle dans le contrôle du pas τ dans l'algorithme.

Théorème 1.2.1 *Soient K un ensemble fermé et convexe dans $I\!R^n$ et F un opérateur de K dans $I\!R^n$ qui est pseudo monotone par rapport à SOL(K,F) sur K et Lipschitz continu de module $L > 0$ sur K. Si $\tau < \frac{1}{L}$, alors la suite $\{x^k\}$ générée par l'algorithme de l'extragradient converge vers une solution de l'VI(K,F).*

Preuve. Soit x^* un élément arbitraire de SOL(K,F). Posons $\rho := 1 - \tau^2 L^2$. Par la condition sur K, nous avons $\rho \in]0, 1[$. D'après le dernier lemme et l'hypothèse sur τ, la suite $\{x^k\}$ est bornée de sorte que, par le théorème de Bolzano-Weierstrass, il y a au moins un point d'adhérence $\bar{x}$ qui doit appartenir à K. Montrons que $\bar{x} \in$ SOL(K,F). Utilisant le dernier lemme, le fait que $\rho \in]0, 1[$ et la positivité de la norme, nous obtenons

$$\rho \sum_{k=0}^{\infty} \|x^k - x^{k+\frac{1}{2}}\|^2 \leq \|x^0 - x^*\|^2 < \infty.$$

Ceci implique que

$$\lim_{k \to \infty} \|x^k - x^{k+\frac{1}{2}}\| = 0,$$

et, par unicité de la limite, que

$$\lim_{k(\in \kappa) \to \infty} \|x^k - x^{k+\frac{1}{2}}\| = 0$$

avec $\{x^k : k \in \kappa\}$ sous-suite convergeant vers $\bar{x}$, c'est-à-dire telle que $\lim_{k(\in \kappa) \to \infty} \|x^k - \bar{x}\| = 0$, ce qui induit que

$$\lim_{k(\in \kappa) \to \infty} x^{k+\frac{1}{2}} = \bar{x}.$$

Par la définition de $x^{k+\frac{1}{2}}$ au Pas 2 de l'algorithme de l'extragradient et par la continuité de F et du projecteur, on voit que

$$\bar{x}$$

$$= \lim_{k(\in \kappa) \to \infty} x^{k+\frac{1}{2}}$$

$$= \lim_{k(\in \kappa) \to \infty} \Pi_K[x^k - \tau F(x^k)]$$

$$= \Pi_K[\bar{x} - \tau F(\bar{x})],$$

ce qui montre que $\bar{x} \in$ SOL(K,F).

Il reste à montrer que la suite $\{x^k\}$ toute entière converge vers $\bar{x}$. Pour cela, il suffit d'appliquer le lemme qui précède avec $x^* = \bar{x}$ pour déduire que la suite positive $\{\|x^k - \bar{x}\|\}$ est monotone décroissante et qu'elle converge par conséquent. Puisque, par unicité de la limite,

$$\lim_{k \to \infty} \|x^k - \bar{x}\| = \lim_{k(\in \kappa) \to \infty} \|x^k - \bar{x}\| = 0,$$

$\{x^k\}$ converge vers $\bar{x}$ comme voulu.

1.3 La méthode de projection sur un hyperplan.

Comme remarqué plus tôt, la méthode de l'extragradient requiert encore la connaissance du module de Lipschitz L de F, ce qui limite son applicabilité puisque de dernier est assez souvent inconnu. Dans cette sous-section, nous présentons une méthode améliorée de type gradient qui ne requiert ni la Lipschitz continuité de F, ni la connaissance a priori de constantes particulières.

Soient K un ensemble fermé et convexe de $\mathbb{R}^n$ et F un opérateur continu de K dans $\mathbb{R}^n$ qui est pseudo monotone sur K par rapport à SOL(K,F). Soit aussi $\tau > 0$ un scalaire fixé. L'algorithme peut être décrit géométriquement comme suit. Soit $x^k \in K$ donné. Tout d'abord, on calcule le point $\Pi_K[x^k - \tau F(x^k)]$. Ensuite, par une simple routine de recherche linéaire de type Armijo, on cherche, sur le segment de droite joignant x^k et $\Pi_K[x^k - \tau F(x^k)]$, le point z^k tel que l'hyperplan

$$H^k := \{x \in \mathbb{R}^n : \langle F(z^k), x - z^k \rangle = 0\}$$

sépare strictement x^k de SOL(K,F). On projette alors x^k sur H^k et le point résultant sur K, ce qui donne x^{k+1}. On peut montrer que x^{k+1} est plus proche de SOL(K,F) que x^k. Si on se réfère à la discussion faite après la présentation de l'algorithme de l'extragradient pour le cas $K = \mathbb{R}^n$, on peut aisément voir quels sont les principaux changements. Tout d'abord, puisqu'on ne peut plus utiliser l'information sur le module de Lipschitz L pour déterminer un pas τ le long de $-F(x^k)$ qui donne un hyperplan séparateur H^k qui convienne, on effectue une recherche linéaire le long de la direction $-F(x^k)$ pour trouver un tel hyperplan. Ensuite, à la place de se déplacer à partir de x^k *le long de* la direction de la projection de x^k sur H^k, on prend x^{k+1} comme étant exactement une telle projection. La méthode qui en résulte requiert trois projections par itération, deux sur K et une sur H^k. Cette dernière projection est facilement effectuée et est donnée par une formule explicite. L'algorithme de projection sur un hyperplan se synthétise alors comme ce qui suit.

Algorithme 1.3.1

Données : $x^0 \in K, \tau > 0$, *et* $\sigma \in]0,1[$.

Pas 0 : *Poser* $k := 0$.

Pas 1 : *Si* $x^k = \Pi_K[x^k - \tau F(x^k)]$, *stop.*

Pas 2 : *Calculer*

$$y^k := \Pi_K[x^k - \tau F(x^k)],$$

et trouver le plus petit entier positif i_k *tel qu'avec* $i = i_k$,

$$\langle F(2^{-i}y^k + [1 - 2^{-i}]x^k), x^k - y^k \rangle \geq \frac{\sigma}{\tau}\|x^k - y^k\|^2.$$

Pas 3 : *Poser*

$$z^k := 2^{-i_k}y^k + (1 - 2^{-i_k})x^k,$$

et

$$w^k := \Pi_{H^k}(x^k) = x^k - \frac{\langle F(z^k), x^k - z^k\rangle}{\|F(z^k)\|^2} F(z^k).$$

Pas 4 : *Poser* $x^{k+1} := \Pi_K(w^k)$ *et* $k := k+1$ *; revenir au Pas 1.*

En notant que $z^k = \tau_k y^k + (1-\tau_k)x^k$, où $\tau_k := 2^{-i_k} \Leftrightarrow \tau_k(x^k - y^k) = x^k - z^k$, nous avons

$$w^k = x^k - \tilde{\tau}_k F(z^k)$$

où

$$\tilde{\tau}_k := \tau_k \frac{\langle F(z^k), x^k - y^k\rangle}{\|F(z^k)\|^2}.$$

D'où

$$x^{k+1} = \Pi_k \left(x^k - \tilde{\tau}_k F[\tau_k y^k + (1-\tau_k)x^k] \right)$$

Cette expression montre, une nouvelle fois, que la méthode de projection sur un hyperplan est étroitement liée à la méthode de l'extragradient. En effet, quand $\tau_k = 1$, on obtient $z^k = y^k$ et

$$x^{k+1} = \Pi_K[x^k - \tilde{\tau}_k F(y^k)];$$

Cette formule d'itération diffère de celle dans la méthode de l'extragradient en ce sens que la longueur de pas variable $\tilde{\tau}_k$ est employée à la place de celle fixée τ dans la seconde projection.

L'itération de l'algorithme de projection sur un hyperplan peut sembler plus compliquée que celles des algorithmes de projection considérés jusqu'ici. Mais la complication est plus superficielle que réelle.

Pour prouver la convergence de l'algorithme de projection sur un hyperplan, nous avons besoin de trois résultats préparatoires. Le premier reformule deux propriétés bien connues du projecteur euclidien afin de faciliter leur emploi lorsque l'on manipule des sous-suites.

Lemme 1.3.1 *Soit K un ensemble non vide, fermé et convexe. Les assertions qui suivent tiennent.*

1. *Pour tout x,y dans $I\!R^n$,*

$$\|\Pi_K(x) - \Pi_K(y)\|^2 \leq \|x-y\|^2 - \|\Pi_K(x) - x + y - \Pi_K(y)\|^2.$$

2. *Pour tout $x \in K$ et pour tout $y \in I\!R^n$,*

$$\langle x - \Pi_K(y), x - y\rangle \geq \|x - \Pi_K(y)\|^2.$$

Preuve.

1. $\|\Pi_K(x) - x + y - \Pi_K(y)\|^2$

   $= \|\Pi_K(x) - \Pi_K(y)\|^2 + \|x-y\|^2 - 2\langle \Pi_K(x) - \Pi_K(y), x - y\rangle$

   $\leq -\|\Pi_K(x) - \Pi_K(y)\|^2 + \|x-y\|^2$ en utilisant la co-coercivité de module 1 du projecteur.

2. découle de la dernière inégalité, puisque $x = \Pi_K(x)$ pour $x \in K$.

Le prochain lemme montre, entre autres, que la recherche linéaire du Pas 2 se termine en temps fini, si bien que la méthode de projection sur un hyperplan est bien définie.

Lemme 1.3.2 *Supposons que F est continu sur K et que x^k est un point de K qui n'est pas solution de l'VI(K,F). Les assertions qui suivent tiennent.*

1. *il existe un entier fini $i_k \geq 0$ qui vérifie l'inégalité définissant la recherche linéaire du Pas 2 de l'algorithme de projection sur un hyperplan ;*
2. $\langle F(z^k), x^k - z^k \rangle > 0.$

Preuve.

1. Supposons par l'absurde que pour tout entier positif i nous ayons
$$\langle F(2^{-i}y^k + [1 - 2^{-i}]x^k), x^k - y^k \rangle < \frac{\sigma}{\tau}\|x^k - y^k\|^2.$$
En multipliant par τ et en passant à la limite $i \to \infty$, on obtient,
$$\tau\langle F(x^k), x^k - y^k \rangle \leq \sigma\|x^k - y^k\|^2.$$
Se rappelant la définition de y^k au Pas 2 et en utilisant un artifice de calcul, puis en se rappelant la seconde assertion du dernier lemme, on obtient
$$\sigma\|x^k - y^k\|^2$$
$$\geq \tau\langle F(x^k), x^k - y^k \rangle$$
$$= \langle x^k - [x^k - \tau F(x^k)], x^k - \Pi_K[x^k - \tau F(x^k)] \rangle$$
$$\geq \|x^k - y^k\|^2.$$
Puisque $\sigma \in]0, 1[$, on a $x^k = y^k = \Pi_K[x^k - \tau F(x^k)]$. Mais ceci contredit l'hypothèse selon laquelle $x^k \notin$ SOL(K,F).
2. $\langle F(z^k), x^k - z^k \rangle = 2^{-i_k}\langle F(z^k), x^k - y^k \rangle \geq 2^{-i_k}\frac{\sigma}{\tau}\|x^k - y^k\|^2 > 0$ car $x^k \neq y^k$ étant donné que $x^k \notin$ SOL(K,F).

Si F est pseudo monotone sur K par rapport à SOL(K,F), la seconde partie du lemme qui précède montre que l'hyperplan H^k sépare strictement x^k de SOL(K,F). En effet, étant donné que la combinaison convexe z^k appartient à K, nous avons, pour toute solution x^* de l'VI(K,F),
$$\langle F(z^k), z^k - x^* \rangle \geq 0 > \langle F(z^k), z^k - x^k \rangle.$$
D'où SOL(K,F) et x^k se trouvent en les côtés opposés de l'hyperplan H^k. Cette propriété est au coeur de l'algorithme de projection sur un hyperplan et montre pourquoi la propriété de pseudo monotonicité de la paire (K,F) est nécessaire.

On en arrive au troisième et dernier résultat préparatoire précédant la preuve du théorème de convergence de l'algorithme de projection sur un hyperplan.

Proposition 1.3.1 *Soient K un ensemble fermé et convexe dans $\mathbb{R}^n$ et F un opérateur continu de K dans $\mathbb{R}^n$ qui est pseudo monotone sur K par rapport à SOL(K,F). Soit aussi la suite $\{x^k\}$ produite par l'algorithme de projection sur un hyperplan. Les quatre assertions qui suivent tiennent.*

1. *Pour toute solution x^* de l'VI(K,F), la suite positive $\{\|x^k - x^*\|\}$ est décroissante et, par conséquent, converge.*
2. *La suite $\{x^k\}$ est bornée.*
3. *$lim_{k\to\infty} \langle F(z^k), x^k - z^k\rangle = 0$.*
4. *Si un point d'adhérence de $\{x^k\}$ est solution de l'VI(K,F), alors le suite toute entière converge vers cette solution.*

Preuve.

1. Soit
$$H^k_\leq := \{x \in \mathbb{R}^n : \langle F(z^k), x - z^k\rangle \leq 0\}.$$
On a, par la discussion faite plus haut,
$$\text{SOL}(K,F) \subseteq H^k_\leq \not\ni x^k.$$
Donc, puisque x^* appartient respectivement à K et $H^k_\leq$,
$$x^* = \Pi_K(x^*) = \Pi_{H^k_\leq}(x^*)$$
alors que w^k est la projection de $x^k(\notin H^k_\leq)$ sur $H^k_\leq$. En conséquent, par la définition de x^{k+1} et la première assertion du premier résultat préparatoire (appliquée deux fois) respectivement, nous avons, pour tout $x^* \in \text{SOL}(K,F)$,

$\|x^{k+1} - x^*\|^2$

$= \|\Pi_K(w^k) - \Pi_K(x^*)\|^2$

$\leq \|w^k - x^*\|^2 - \|\Pi_K(w^k) - w^k + x^* - \Pi_K(x^*)\|^2$

$= \|w^k - x^*\|^2 - \|\Pi_K(w^k) - w^k\|^2$ puisque $x^* \in K$

$= \|\Pi_{H^k_\leq}(x^k) - \Pi_{H^k_\leq}(x^*)\|^2 - \|\Pi_K(w^k) - w^k\|^2$ par la discussion faite plus tôt et puisque $x^* \in H^k_\leq$

$\leq \|x^k - x^*\|^2 - \|\Pi_{H^k_\leq}(x^k) - x^k + x^* - \Pi_{H^k_\leq}(x^*)\|^2 - \|\Pi_K(w^k) - w^k\|^2$

$= \|x^k - x^*\|^2 - \|w^k - x^k\|^2 - \|\Pi_K(w^k) - w^k\|^2$

$\leq \|x^k - x^*\|^2$ par positivité de la norme.

2. Découle directement de l'assertion qui précède.
3. En utilisant la pénultième inégalité dans la dernière chaîne d'inégalités, on voit que $0 \leq \|w^k - x^k\|^2 \leq \|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2 - \|\Pi_K(w^k) -$

$w^k\|^2 \leq \|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2$.

En utilisant la première assertion et le théorème de l'étau, puis la définition analytique de w^k, on déduit

$$0 = \lim_{k\to\infty} \|x^k - w^k\| = \lim_{k\to\infty} \frac{\langle F(z^k), x^k - z^k\rangle}{\|F(z^k)\|}.$$

La seconde assertion implique clairement le caractère borné de $\{z^k\}$ et donc, par continuité de F, celui de $\{\|F(z^k)\|\}$. D'où, on obtient

$$\lim_{k\to\infty} \langle F(z^k), x^k - z^k\rangle = 0.$$

4. Soient les deux sous-suites $\{x^{q_k}\}$ et $\{x^{n_k}\}$ de $\{x^k\}$ tendant respectivement vers les deux points d'adhérence x^* et x^{**} solutions de l'VI(K,F), ce qui implique que $\lim_{q_k\to\infty} \|x^{q_k} - x^*\| = \lim_{n_k\to\infty} \|x^{n_k} - x^{**}\| = 0$. On doit prouver que $x^* = x^{**}$. Or, par la première assertion, les suites $\{\|x^k - x^*\|\}$ et $\{\|x^k - x^{**}\|\}$ sont convergentes. Dès lors, par unicité de la limite, $\lim_{k\to\infty} \|x^k - x^*\| = \lim_{k\to\infty} \|x^k - x^{**}\| = 0$.

   Ainsi, en appliquant l'inégalité triangulaire et le théorème de l'étau à la relation $0 \leq \|x^* - x^{**}\| \leq \|x^* - x^k\| + \|x^k - x^{**}\|$, on déduit que $x^* = x^{**}$.

Nous sommes maintenant prêts à établir et prouver le résultat promis de convergence de l'algorithme de convergence sur un hyperplan.

Théorème 1.3.1 *Soient K un ensemble fermé et convexe dans $I\!R^n$ et F un opérateur continu de K dans $I\!R^n$ qui est pseudo monotone sur K par rapport à SOL(K,F). Si $\{x^k\}$ est une suite produite par l'algorithme de projection sur un hyperplan, alors $\{x^k\}$ converge vers une solution de l'VI(K,F).*

Posons, pour alléger les notations, $t_k := 2^{-i_k}$. Par la troisième assertion de la dernière proposition,

$$\lim_{k\to\infty} \langle F(z^k), x^k - z^k\rangle = 0 = \lim_{k\to\infty} t_k\langle F(z^k), x^k - y^k\rangle.$$

On considère deux cas. Cas 1 : $\lim\sup_{k\to\infty} t_k > 0$, c'est-à-dire qu'il existe $\bar{t} > 0$ et une sous-suite $\{x^k : k \in \kappa\}$ telle que $t_k \geq \bar{t}$ pour tout $k \in \kappa$. Utilisant la dernière limite, on obtient alors, par unicité de la limite,

$$\lim_{k(\in\kappa)\to\infty} \langle F(z^k), x^k - y^k\rangle = 0.$$

Cette dernière limite implique, par application du théorème de l'étau à l'inégalité définissant la recherche linéaire de l'algorithme de projection sur un hyperplan et par définition de z^k, que

$$\lim_{k(\in\kappa)\to\infty} \|x^k - y^k\| = 0 = \lim_{k(\in\kappa)\to\infty} \|x^k - \Pi_K[x^k - \tau F(x^k)]\|.$$

Puisque $\{x^k : k \in \kappa\}$ est bornée par la seconde assertion de la dernière proposition, on peut supposer, par passage au théorème de Bolzano-Weierstrass et donc à une sous suite de $\{x^k : k \in \kappa\}$ et à son point d'adhérence x^∞,

$\lim_{k(\in\kappa)\to\infty} x^k = x^\infty$ où la nouvelle sous-suite fille est, sans perte de généralité, indexée comme la sous-suite mère. D'où, on obtient, par continuité de F,

$$\|x^\infty - \Pi_K[x^\infty - \tau F(x^\infty)]\| = 0.$$

Ceci montre que x^∞ est une solution de l'VI(K,F). Par la dernière assertion de la dernière proposition, on conclut par conséquent que la suite $\{x^k\}$ toute entière converge vers x^∞, ce qui achève la preuve dans ce premier cas.

Cas 2 : $\lim_{k\to\infty} t_k = 0$. Posons

$$\bar{z}^k := 2t_k y^k + (1-2t_k)x^k.$$

Comme précédemment, nous pouvons supposer, sans perte de généralité, qu'une certaine sous-suite $\{x^k : k \in \kappa\}$ converge vers une limite x^∞. Puisque, par unicité de la limite, la sous-suite $\{t_k : k \in \kappa\}$ converge vers zéro, nous avons

$$\lim_{k(\in\kappa)\to\infty} \bar{z}^k = x^\infty.$$

Par les règles du Pas 2 de l'algorithme de projection sur un hyperplan, et étant donné que $\{t_k\} \to 0$, c'est-à-dire que $\{i_k\} \to \infty$, nous avons aussi

$$\langle F(\bar{z}^k), x^k - y^k\rangle < \frac{\sigma}{\tau}\|x^k - y^k\|^2.$$

Passant à la limite $k(\in \kappa) \to \infty$, nous obtenons, par continuité du projecteur et de F,

$$\langle F(x^\infty), x^\infty - y^\infty\rangle \le \frac{\sigma}{\tau}\|x^\infty - y^\infty\|^2,$$

où $y^\infty := \Pi_K[x^\infty - \tau F(x^\infty)]$. Par la seconde assertion du premier résultat préparatoire, il s'ensuit, après multiplication par τ, que

$\sigma\|x^\infty - y^\infty\|^2$

$\ge \tau\langle F(x^\infty), x^\infty - y^\infty\rangle$

$= \langle x^\infty - [x^\infty - \tau F(x^\infty)], x^\infty - \Pi_K[x^\infty - \tau F(x^\infty)]\rangle$ par utilisation d'un artifice de calcul et définition de y^∞

$\ge \|x^\infty - \Pi_K[x^\infty - \tau F(x^\infty)]\|^2$

$= \|x^\infty - y^\infty\|^2$ par définition de y^∞.

Puisque $\sigma \in]0,1[$, cette dernière inégalité implique que $x^\infty = y^\infty = \Pi_K[x^\infty - \tau F(x^\infty)]$ de sorte que x^∞ est une solution de l'VI(K,F). La suite $\{x^k\}$ toute entière converge alors vers x^∞ par la dernière assertion de la dernière proposition.

Il existe une variante simple de l'algorithme de projection sur un hyperplan qui peut être digne d'intérêt. On note qu'au Pas 3 de cet algorithme, on obtient x^{k+1} en projetant tout d'abord x^k sur H^k, ce qui donne w^k, et en projetant ensuite w^k sur K. Il est possible de montrer, par une simple modification des preuves, que le dernier théorème tient toujours si l'on calcule x^{k+1} comme étant la projection de x^k sur $H^k_\leq \cap K$. Des considérations géométriques évidentes montre que, étant donné x^k, le point x^{k+1} calculé selon cette procédure modifiée, est plus proche de SOL(K,F) que le point x^{k+1} calculé comme dans le Pas 3 de l'algorithme de projection sur un hyperplan. Ceci suggère que la variante puisse être plus rapide que la méthode originale. Cet avantage doit être tempéré par le fait que projeter sur $H^k_\leq \cap K$ peut être plus difficile que projeter sur $H^k_\leq$ et K séparément. Par exemple, supposons que K est un rectangle ou une sphère. Les projections sur H^k et K peuvent être calculées analytiquement. Cependant, la projection sur $H^k_\leq \cap K$ se substitue à la solution d'un problème de minimisation convexe non trivial.

Chapitre 2

Les problèmes d'équilibre.

2.1 Préambule.

L'analyse d'un problème général d'équilibre, noté EP,

$$\text{trouver } x^* \in K \text{ tel que } f(x^*, y) \geq 0 \ \forall y \in K,$$

où $f : K \times K \to \mathbb{R}$, avec $f(x,x) = 0 \ \forall x \in K$ un sous-ensemble de l'espace réflexif de Banach X, a mené à une approche unifiée dans le développement de l'étude de différents sujets en optimisation tels que des problèmes d'extremum sous contraintes et des inéquations variationnelles. En particulier, ces dernières ont été extensivement étudiées puisqu'elles permettent de généraliser les conditions classiques d'optimalité pour des problèmes d'extremum sous contraintes.

Si nous définissons $f(x,y) := \langle F(x), y - x \rangle$, alors le EP prend la forme de l'inéquation variationnelle classique, notée VI,

$$\text{trouver } x^* \in K \text{ tel que } \langle F(x^*), y - x^* \rangle \geq 0 \ \forall y \in K,$$

où $F : K \to X^*$, $K \subseteq X$ et $\langle \cdot, \cdot \rangle$ est le produit scalaire entre X et son dual topologique X^*.

Si $f(x,y) := h(y) - h(x)$, alors le EP est équivalent au problème d'optimisation

$$\begin{aligned} &\text{trouver } x^* \in K \text{ tel que } h(y) - h(x^*) \geq 0 \ \forall y \in K, \\ \Leftrightarrow \ &\text{trouver } x^* \in K \text{ tel que } h(y) \geq h(x^*) \ \forall y \in K, \\ \Leftrightarrow \ &\min\nolimits_{y \in K} h(y), \end{aligned}$$

où $h : K \to \mathbb{R}$.

Le but de cette section est de montrer qu'il est possible, en exploitant une formulation en point fixe du EP qui convienne, de définir une classe de méthodes itératives pour résoudre des VI et des problèmes d'extremum sous contraintes.

Dans la prochaine section, nous introduirons une régularisation du EP (adaptée à notre cadre de travail) à laquelle nous nous référerons comme étant le problème d'équilibre auxiliaire ; celui-ci est un problème d'équilibre généralisé qui est équivalent au EP d'origine.

Dans la section suivante, au moyen de la formulation en point fixe du problème d'équilibre auxiliaire, nous montrerons qu'il est possible de définir une suite fortement convergente vers une solution de ce dernier et donc du EP.

Et dans la dernière section, nous présenterons certaines applications aux problèmes d'extremum sous contraintes.

2.2 Le problème d'équilibre auxiliaire.

La plupart des algorithmes développés pour résoudre le EP peuvent être dérivés de formulations équivalentes de celui-ci. Nous allons focaliser notre attention sur des formulations en point-fixe : nous montrerons que de telles formulations mènent à une généralisation de méthodes développées pour des inéquations variationnelles et des problèmes d'optimisation.

Le résultat préliminaire qui suit établit la formulation équivalente du EP mentionnée ci-avant.

Lemme 2.2.1 *Supposons que $f(x,x) = 0 \ \ \forall x \in K$. Alors, les assertions suivantes sont équivalentes :*

1. *$\exists x^* \in K$ tel que $f(x^*, y) \geq 0 \ \forall y \in K$.*
2. *$x^* \in K$ est une solution du problème de $min_{y\in K} \ f(x^*, y)$.*

Preuve *$x^* \in K$ est une solution du problème de $min_{y\in K} \ f(x^*, y)$*

$\Leftrightarrow f(x^*, y) \geq f(x^*, x^*) = 0 \ \forall y \in K$

$\Leftrightarrow \exists x^* \in K$ *tel que* $f(x^*, y) \geq 0 \ \forall y \in K$.

Si nous supposons que, $\forall x^* \in K$, le problème de $min_{y\in K} \ f(x^*, y)$ a une unique solution, nous pouvons définir la méthode itérative générale suivante (qui porte le nom d'*algorithme général*) :

Algorithme 2.2.1

1. *Poser $k := 0$ et choisir $x^0 \in K$;*
2. *Calculer x^{k+1} la solution du problème de $min_{y\in K} \ f(x^k, y)$;*
3. *Si $\|x^{k+1} - x^k\| < \mu$ pour un certain $\mu > 0$ fixé, alors STOP.*

   Sinon, poser $k := k + 1$ et revenir au second pas.

Malheureusement, dans la plupart des cas, il n'est pas possible (ou, du moins, pas pratique) d'appliquer directement cet agorithme au EP. Il est nécessaire d'introduire un problème d'équilibre auxiliaire, équivalent à celui donné, pour lequel la procédure décrite ci-avant mène à une solution du EP.

Proposition 2.2.1 *Soient un ensemble convexe K, une fonction $f(x,y)$, avec $f(x,x) = 0\ \forall x \in K$, telle que, pour $x^* \in K$, $f(x^*,\cdot)$ est convexe et différentiable, et une constante $\epsilon > 0$. Supposons que $H(x,y) : K \times K \to I\!R$ est positive, telle que, pour $x \in K$, $H(x,\cdot)$ est différentiable, et telle que*

1. $H(x,x) = 0\ \forall x \in K$;
2. $H'_y(x,x) = 0\ \forall x \in K$.

Alors x^* est solution du EP si et seulement s'il est solution du problème d'équilibre auxiliaire, noté AEP,

$$\text{trouver } x^* \in K \text{ tel que } \epsilon f(x^*,y) + H(x^*,y) \geq 0\ \forall y \in K.$$

Preuve Il est évident (puisque H est positive) que si x^* est solution du EP, alors il est aussi solution du AEP.

Inversément, supposons que x^* est solution du AEP. Or, en appliquant le lemme vu en début de section, on peut affirmer que x^* solutionne ce dernier problème si et seulement si x^* est un point minimum du problème

$$min_{y\in K}[\epsilon f(x^*,y) + H(x^*,y)].$$

Puisque K est convexe, en utilisant la condition nécessaire et suffisante d'optimalité en programmation convexe, x^* est une solution optimale de ce problème si et seulement si

$0 \in [\epsilon f'_y(x^*,x^*) + H'_y(x^*,x^*)] + N_K(x^*)$, avec $N_K(x^*) := \{w : \langle w, y - x^*\rangle \leq 0\ \forall y \in K\}$

$\Leftrightarrow [-\epsilon f'_y(x^*,x^*) - H'_y(x^*,x^*)] \in N_K(x^*)$, avec $N_K(x^*) := \{w : \langle w, y - x^*\rangle \leq 0\ \forall y \in K\}$

$\Leftrightarrow \langle -\epsilon f'_y(x^*,x^*) - H'_y(x^*,x^*), y - x^*\rangle \leq 0\ \forall y \in K$

si bien que

$$\langle \epsilon f'_y(x^*,x^*), y - x^*\rangle \geq 0\ \forall y \in K$$

puisque, par hypothèse, $H'_y(x^*,x^*) = 0$ Divisant par ϵ, nous obtenons que cette dernière inéquation implique, par la convexité de $f(x^*,.)$, que

$$f(x^*,y) \geq f(x^*,x^*) = 0\ \forall y \in K.$$

Nous avons démontré au passage que

Corollaire 2.2.1 *x^* est solution du EP si et seulement s'il est une solution optimale du problème d'extremum*

$$min_{y\in K}[\epsilon f(x^*,y) + H(x^*,y)].$$

Dans la prochaine section, nous verrons qu'en appliquant l'algorithme général au AEP pour un choix judicieux de la fonction H, il sera possible de définir une suite x^k qui converge vers une solution du EP.

2.3 Le principe du problème auxiliaire.

Venons-en à l'extension proprement dite du principe du problème auxiliaire au EP.

Supposons que l'on dispose d'une fonction $G : K \to \mathbb{R}$ fortement convexe et différentiable, et d'une constante $\epsilon > 0$. Si nous posons $H(x, y) := G(y) - G(x) - \langle G'(x), y - x \rangle$ (≥ 0 puisque G est fortement convexe), le AEP qui en résulte (et qui, comme nous le verrons par après, jouit de propriétés très intéressantes) prend alors la forme :

trouver $x^* \in K$ tel que $\epsilon f(x^*, y) - \langle G'(x^*), y - x^* \rangle + G(y) - G(x^*) \geq 0 \; \forall y \in K$.

Lemme 2.3.1 *Soient un ensemble convexe K et une fonction $f(x, y)$, avec $f(x, x) = 0 \; \forall x \in K$, telle que, pour $x^* \in K$, $f(x^*, \cdot)$ est convexe et différentiable. Alors, x^* est solution du EP si et seulement s'il est solution du AEP qui précède.*

Preuve Découle directement de la proposition vue à la section précédente en notant que la fonction $H(x, y) := G(y) - G(x) - \langle G'(x), y - x \rangle$ vérifie les hypotèses requises : en effet, elle est positive (puisque G est fortement convexe) et

1. $H(x, x) := G(x) - G(x) - \langle G'(x), x - x \rangle = 0 \; \forall x \in K$
2. $H'_y(x, x) := G'(x) - 0 - G'(x) = 0 \; \forall x \in K$.

Prenant en compte le corollaire vu à la section précédente, nous avons que x^* est solution du EP si et seulement s'il est solution optimale du problème d'extremum (donc inchangé à une constante près)

$$min_{y \in K} [\epsilon f(x^*, y) - \langle G'(x^*), y \rangle + G(y)].$$

Appliquant l'algorithme général au AEP, nous obtenons la méthode itérative qui suit :

Algorithme 2.3.1

1. *Poser $k := 0$ et choisir $x^0 \in K$;*
2. *Calculer x^{k+1} la solution du problème $P(k)$:*

$$min_{y \in K} [\epsilon f(x^k, y) - \langle G'(x^k), y \rangle + G(y)];$$

3. *Si $\|x^{k+1} - x^k\| < \mu$ pour un certain $\mu > 0$ fixé, alors STOP.*

   Sinon, poser $k := k + 1$ et revenir au second pas.

Remarque 2.3.1 *Nous observons que $P(k)$ a une unique solution si $f(x^k, \cdot)$ est convexe, puisque $\langle G'(x^k), \cdot \rangle$ est linéaire et $G(\cdot)$ est fortement convexe.*

Théorème 2.3.1 *Supposons que les conditions suivantes sont vérifiées :*

1. *$f(x, \cdot)$ est une fonction convexe et semi-continue inférieurement $\forall x \in K$;*
2. *$f(\cdot, y)$ est une fonction continue , sur tout sous-espace de X de dimension finie, $\forall y \in K$;*
3. *f est fortement monotone de module a sur K ;*
4. *G est fortement convexe de module b sur K ;*
5. *il existe des constantes $\alpha, \beta > 0$ telles que $\forall x, y, z \in K$:*

$$f(y, x) + f(z, y) \geq f(z, x) - \alpha\|x - y\|^2 - \beta\|y - z\|^2.$$

Alors, si $\epsilon \leq \frac{b}{2\alpha}$ et $\beta < a$, la suite x^k, définie dans l'algorithme qui précède, converge fortement vers la solution x^* du EP.

Preuve Au vu des hypothèses, on peut affirmer que la solution x^* du EP existe et est unique(voir Theorem 1 dans [2]), et qu'il en va de même pour la solution de $P(k)$.

Considérons la fonction positive

$\Lambda(y) := G(x^*) - G(y) - \langle G'(y), x^* - y\rangle \geq \frac{b}{2}\|x^* - y\|^2 \geq 0$ (puisque G est fortement convexe), et la différence

$\Lambda(x^k) - \Lambda(x^{k+1})$

$:= G(x^{k+1}) - G(x^k) - \langle G'(x^k), x^* - x^k\rangle + \langle G'(x^{k+1}), x^* - x^{k+1}\rangle$

$+\langle G'(x^k), x^{k+1}\rangle - \langle G'(x^k), x^{k+1}\rangle$

$= G(x^{k+1}) - G(x^k) - \langle G'(x^k), x^{k+1} - x^k\rangle + \langle G'(x^{k+1}) - G'(x^k), x^* - x^{k+1}\rangle$

$\geq \frac{b}{2}\|x^{k+1} - x^k\|^2 + \epsilon[f(x^k, x^{k+1}) - f(x^k, x^*)]$

où l'inégalité qui précède est due à la convexité forte de G et au fait que, puisque x^{k+1} résout $P(k)$, nous avons la relation $\langle \epsilon f'_y(x^k, x^{k+1}) - G'(x^k) + G'(x^{k+1}), y - x^{k+1}\rangle \geq 0 \ \forall y \in K$ (en utilisant la condition nécessaire et suffisante d'optimalité en programmation convexe), qui, pour $y := x^*$, implique $\langle G'(x^{k+1}) - G'(x^k), x^* - x^{k+1}\rangle \geq \langle \epsilon f'_y(x^k, x^{k+1}), x^{k+1} - x^*\rangle \geq \epsilon[f(x^k, x^{k+1}) - f(x^k, x^*)]$ étant donné que $f(x^k, \cdot)$ est convexe.

Par conséquent, en utilisant un artifice de calcul,

$\Lambda(x^k) - \Lambda(x^{k+1}) \geq \frac{b}{2}\|x^{k+1} - x^k\|^2 + \epsilon[f(x^k, x^{k+1}) + f(x^*, x^k)] - \epsilon[f(x^k, x^*) + f(x^*, x^k)]$.

En exploitant les cinquième et troisième hypothèses, puis, respectivement, les inégalités $f(x^*, x^{k+1}) \geq 0$ (vu que x^* est solution du EP), $\epsilon \leq \frac{b}{2\alpha}$ et $\beta < a$, nous obtenons

$\Lambda(x^k) - \Lambda(x^{k+1}) \geq \frac{b}{2}\|x^{k+1} - x^k\|^2 + \epsilon[f(x^*, x^{k+1}) - \alpha\|x^{k+1} - x^k\|^2 - \beta\|x^k - x^*\|^2] + \epsilon a\|x^k - x^*\|^2 \geq (\frac{b}{2} - \epsilon\alpha)\|x^{k+1} - x^k\|^2 + \epsilon(a - \beta)\|x^k - x^*\|^2 \geq \epsilon(a-\beta)\|x^k - x^*\|^2 \geq 0$.

Dès lors, la suite $\Lambda(x^k)$ est décroissante et positive, et donc convergente si bien que $\Lambda(x^k) \to \bar{\Lambda}$ quand $k \to \infty$.

Ainsi, en appliquant le théorème de l'étau aux inégalités précédentes, $x^k \to x^*$ quand $k \to \infty$.

2.4 Applications aux problèmes d'optimisation.

Considérons l'inéquation variationnelle généralisée, notée GVI,

$$\text{trouver } x^* \in K \text{ tel que } \langle F(x^*), y - x^*\rangle \geq \phi(x^*) - \phi(x) \ \forall y \in K,$$

où $F : K \to X^*$ et $\phi : K \to \mathbb{R}$ est une fonction convexe propre semi-continue inférieurement. Si on pose

$$f(x,y) := \langle F(x), y - x\rangle - \phi(x) + \phi(y),$$

alors l'GVI est équivalente au EP de sorte que l'on peut appliquer l'algorithme vu à la section précédente. On obtient alors (en utilisant le fait qu'une minimisation se fait à une constante près) :

Algorithme 2.4.1

1. *Poser* $k := 0$ *et choisir* $x^0 \in K$;
2. *Calculer* x^{k+1} *la solution du problème :*

$$min_{y\in K}[\langle \epsilon F(x^k) - G'(x^k), y\rangle + G(y) + \epsilon\phi(y)];$$

3. *Si* $\|x^{k+1} - x^k\| < \mu$ *pour un certain* μ *fixé, alors STOP.*

   Sinon, poser $k := k+1$ *et revenir au second pas.*

Théorème 2.4.1 *Supposons que les conditions suivantes sont vérifiées :*

1. *F est un opérateur fortement monotone de module a sur K qui est continu sur tout sous-espace de X de dimension finie ;*
2. *G est fortement convexe de module b sur K ;*
3. *il existe des constantes* $\alpha, \beta > 0$ *telles que* $\forall x, y, z \in K$:

$$\langle F(z) - F(y), y - x\rangle \geq -\alpha\|x - y\|^2 - \beta\|y - z\|^2$$

Alors, si $\epsilon \leq \frac{b}{2\alpha}$ *et* $\beta < a$*, la suite* x^k*, définie dans l'algorithme qui précède, converge fortement vers la solution* x^* *de l'GVI.*

Preuve Découle du théorème vu à la section précédente, en posant

$$f(x,y) := \langle F(x), y-x\rangle - \phi(x) + \phi(y).$$

En effet,

1. $f(x,\cdot)$ est semi-continue inférieurement (évidemment) et convexe $\forall x \in K$ puisque $\forall y_1, y_2 \in K$:

   $f(x,y_2) := \langle F(x), y_2 - x\rangle - \phi(x) + \phi(y_2) \geq [\langle F(x), y_1 - x\rangle - \phi(x) + \phi(y_1)] + \langle F(x) + \phi'(y_1), y_2 - y_1\rangle =: f(x,y_1) + \langle f'_y(x,y_1), y_2 - y_1\rangle$

   (étant donné que ϕ est convexe et qu'on a donc $\phi[y_2] \geq \phi[y_1] + \langle \phi'[y_1], y_2 - y_1\rangle$) ;
2. $f(\cdot, y)$ est évidemment continue, sur tout sous-espace de X de dimension finie, $\forall y \in K$;
3. f est fortement monotone de module a sur K puisque $\forall x, y \in K$:

   $f(x,y) + f(y,x) := [\langle F(x), y-x\rangle - \phi(x) + \phi(y)] + [\langle F(y), x-y\rangle - \phi(y) + \phi(x)] \leq -a\|y-x\|^2$

   (étant donné que F est fortement monotone et qu'on a donc $\langle F[y] - F[x], y-x\rangle \geq a\|y-x\|^2$) ;
4. G est fortement convexe de module b sur K par hypothèse ;
5. il existe des constantes $\alpha, \beta > 0$ telles que $\forall x, y, z \in K$, on a

   $f(y,x) + f(z,y) - f(z,x) \geq -\alpha\|x-y\|^2 - \beta\|y-z\|^2$

   puique, $\forall x, y, z \in K$,

   $f(y,x) + f(z,y) - f(z,x) := [\langle F(y), x-y\rangle - \phi(y) + \phi(x)] + [\langle F(z), y-z\rangle - \phi(z) + \phi(y)] - [\langle F(z), x-z\rangle + \phi(z) - \phi(x)] \geq -\alpha\|x-y\|^2 - \beta\|y-z\|^2$ (étant donné que, par hypothèse, $\langle F(z) - F(y), y-x\rangle \geq -\alpha\|x-y\|^2 - \beta\|y-z\|^2$).

Montrons maintenant que si F est Lipschitz continu sur K, alors la troisième hypothèse du précédent théorème est satisfaite pour un judicieux choix des constantes α et β.

Proposition 2.4.1 *Supposons que F est Lipschitz continu de module L sur K, alors la troisième hypothèse du précédent théorème est satisfaite à condition que $\sqrt{\alpha\beta} \geq \frac{L}{2}$.*

Preuve Appliquant respectivement l'inégalité de Cauchy-Schwarz, la Lipschitz continuité de F et le fait que $(\sqrt{\alpha}\|x-y\| - \sqrt{\beta}\|y-z\|)^2 \geq 0$, nous avons, $\forall x, y, z \in K$:

$\langle F(z) - F(y), y-x\rangle \geq -\|F(y) - F(z)\|\|x-y\| \geq -L\|y-z\|\|x-y\| \geq -2\sqrt{\alpha\beta}\|y-z\|\|x-y\| \geq -\alpha\|x-y\|^2 - \beta\|y-z\|^2$.

Remarque 2.4.1 *Pour appliquer le théorème qui précède, il suffit de choisir α et β qui vérifient le système qui suit :*

$$\sqrt{\alpha\beta} \geq \frac{L}{2}, \alpha \leq \frac{b}{2\epsilon}, \beta < a, \alpha > 0, \beta > 0.$$

Avant d'en venir aux problèmes d'extremum sous contraintes, introduisons deux définitions qui nous seront utiles.

Définition 2.4.1 *Une fonction $G : X \to I\!R$ est dite différentiable au sens de Gâteaux au point $x^* \in X$ si et seulement si la limite*

$$lim_{t\to 0}\frac{G(x^* + ty) - G(x^*)}{t} =: \langle G'(x^*), y\rangle,$$

avec $G'(x^*) \in X^*$, existe et est finie.

Définition 2.4.2 *Une fonction $h : X \to I\!R$ est dite directionnellement différentiable au sens de Dini au point $x^* \in X$ dans la direction y si et seulement si la limite*

$$lim_{t\to 0^+}\frac{h(x^* + ty) - h(x^*)}{t} =: h'(x^*; y)$$

existe et est finie.

L'analyse de l'GVI nous permet de définir un algorithme pour le problème d'extremum sous contrainte, noté (P), de la forme :

$$min_{y\in K}[\psi(y) + \phi(y)]$$

où $\psi : K \to \mathbb{R}$ est une fonction fortement convexe et différentiable au sens de Gâteaux sur K et $\phi : K \to \mathbb{R}$ est une fonction convexe directionnellement différentiable au sens de Dini sur K en toute direction de K.

Comme mentionné dans la première section, (P) est équivalent au problème d'équilibre

$$\text{trouver } x^* \in K \text{ tel que } f(x^*, y) \geq 0 \ \forall y \in K$$

avec $f(x, y) = \psi(y) + \phi(y) - \psi(x) - \phi(x)$. Nous observons que la troisième condition du théorème vu dans la troisième section ne peut être satisfaite, quelles que ψ et ϕ puissent être (en effet, $\forall x, y \in K$ tels que $x \neq y$: $f(x, y) + f(y, x) := [\psi(y) + \phi(y) - \psi(x) - \phi(x)] + [\psi(x) + \phi(x) - \psi(y) - \phi(y)] = 0 > -a\|x - y\|^2 \ \forall a > 0$). Un problème d'équilibre équivalent peut être formulé comme une condition d'optimalité du premier ordre pour (P).

Proposition 2.4.2 *x^* est une solution de (P) si et seulement s'il est une solution de*

$$\langle\psi'(x^*), y - x^*\rangle + \phi(y) - \phi(x^*) \geq 0 \ \forall y \in K.$$

Preuve Nous observons, en utilisant la condition nécessaire et suffisante d'optimalité en programmation convexe, que x^* est une solution de (P) si et seulement si

$$\langle\psi'(x^*), y - x^*\rangle + \phi'(x^*; y - x^*) \geq 0 \ \forall y \in K$$

(puisque ψ est différentiable au sens de Gâteaux sur K et ϕ est directionnellement différentiable au sens de Dini sur K en toute direction de K). Or, ϕ étant convexe, nous avons (voir [7])

$$\phi(y) - \phi(x^*) \geq \phi'(x^*; y - x^*) \ \forall y \in K,$$

si bien que $\langle \psi'(x^*), y - x^* \rangle \geq -\phi'(x^*; y - x^*) \geq \phi(x^*) - \phi(y) \ \forall y \in K$. Inversément, si $\langle \psi'(x^*), y - x^* \rangle + \phi(y) - \phi(x^*) \geq 0 \ \forall y \in K$ tient, alors, par application du lemme vu à la seconde section, x^* est un point minimum du problème

$$min_{y \in K}[\langle \psi'(x^*), y - x^* \rangle + \phi(y) - \phi(x^*)],$$

dont la condition d'optimalité du premier ordre est donnée par $\langle \psi'(x^*), y - x^* \rangle + \phi'(x^*; y - x^*) \geq 0 \ \forall y \in K$ (en utilisant de nouveau la différentiabilité au sens de Gâteaux de ψ et la différentiabilité directionnelle au sens de Dini de ϕ).

L'algorithme présenté en début de section permet alors de résoudre l'inéquation variationnelle définie dans la proposition qui précède.

2.5 Synthèse.

Nous avons considéré un algorithme de point fixe pour résoudre un problème d'équilibre général. Nous avons alors montré que cette méthode, qui permet de traiter directement le problème d'origine, doit être appliquée à un problème d'équilibre auxiliaire équivalent afin d'obtenir la convergence de l'algorithme qui découle de celle-ci.

Chapitre 3

Extension de l'extragradient.

Comme annoncé, nous allons utiliser la technique du problème auxiliaire pour développer des algorithmes itératifs résolvant des problèmes d'équilibre. Le premier de ces algorithmes est une extension de l'algorithme de l'extragradient aux problèmes d'équilibre. Dans cet algorithme, la bifonction d'équilibre ne doit pas satisfaire une quelconque propriété de monotonicité, mais elle doit satisfaire une certaine condition de type Lipschitz. Pour éviter cette exigence, nous proposerons des recherches linéaires communément utilisées dans les inéquations variationnelles pour obtenir des algorithmes de type projection résolvant des problèmes d'équilibre. On discutera d'applications aux inéquations variationnelles mixtes. On examinera une classe spéciale et certains résultats préliminaires d'un exemple numérique seront présentés.

3.1 Introduction et description du problème.

Soient K un sous-ensemble non vide, fermé et convexe de l'espace euclidien n-dimensionnel $\mathbb{R}^n$, et $f : K \times K \to \mathbb{R} \cup \{\infty\}$. Considérons le problème d'équilibre primal, noté (PEP), qui suit

$$\text{trouver } x^* \in K \text{ tel que } f(x^*, y) \geq 0 \text{ pour tout } y \in K$$

où $f(x, x) = 0$ pour tout $x \in K$. Comme d'usage, on appellera une bifonction satisfaisant cette propriété une *bifonction d'équilibre* sur K.

Les problèmes d'équilibre ont été longuement étudiés en optimisation. Il est ainsi bien connu que de nombreuses classes de problèmes de programmation mathématiques, d'inéquations variationnelles, de théorèmes du point fixe, d'équilibre de Nash dans des jeux non coopératifs et de problèmes minimax, peuvent être formulées comme un (PEP).

Dans ce chapitre, nous utiliserons le principe du problème auxiliaire pour étendre la méthode de l'extragradient au problème d'équilibre (PEP). De cette manière, nous obtiendrons des algorithmes de l'extragradient pour résoudre le problème (PEP). La convergence des algorithmes proposés ne requiert pas que f satisfasse un quelconque type de monotonicité, mais f doit satisfaire une certaine condition de type Lipschitz. Afin d'éviter cette exigence, nous utiliserons la technique de recherche linéaire pour obtenir des algorithmes résolvant (PEP) qui convergent.

Le reste de ce chapitre est organisé comme suit. Dans la prochaine section, on donne des formulations en point fixe du problème (PEP). On utilise ensuite ces formulations dans la troisième section pour décrire un algorithme de l'extragradient pour le (PEP). La quatrième section est consacrée à la présentation d'algorithmes de recherche linéaire et de leurs résultats de convergence ne requerrant pas la condition de type Lipschitz. Dans la cinquième section, on discutera d'applications des algorithmes présentés aux inéquations variationnelles mixtes multivaluées. La dernière section contient certains résultats de calculs préliminaires et d'essais.

3.2 Les formulations en point fixe.

Reformulons tout d'abord, dans le contexte qui nous intéresse par la suite, les quelques définitions, bien connues en optimisation, déjà présentées en début du précédent chapitre.

Définition 3.2.1 *Soient K un sous-ensemble non vide et convexe dans $I\!R^n$, et $f : K \times K \to I\!R \cup \{\infty\}$. La bifonction f est dite*

1. *fortement monotone de module $\mu > 0$ si et seulement si $\forall x, y \in K$, nous avons*
$$f(x,y) + f(y,x) \leq -\mu \|x - y\|^2;$$
2. *strictement monotone si et seulement si $\forall x, y \in K$ distincts, nous avons*
$$f(x,y) + f(y,x) < 0;$$
3. *monotone si et seulement si $\forall x, y \in K$, nous avons*
$$f(x,y) + f(y,x) \leq 0;$$
4. *pseudo monotone si et seulement si $\forall x, y \in K$, nous avons*
$$f(x,y) \geq 0 \text{ implique } f(y,x) \leq 0;$$
5. *Lipschitz continue de module L si et seulement si $\forall x, y \in K$, nous avons*
$$\|f(x) - f(y)\| \leq L\|x - y\|.$$

Par les définitions précitées, nous avons évidemment les implications qui suivent :

monotonicité forte $\Rightarrow$ monotonicité stricte $\Rightarrow$ monotonicité $\Rightarrow$ pseudo monotonicité.

Nous considérons aussi le problème d'équilibre dual du (PEP), noté (DEP), qui suit :

$$\text{trouver } x^* \in K \text{ tel que } f(y, x^*) \leq 0 \text{ pour tout } y \in K.$$

Pour tout $x \in K$, soit

$$L_f(x) := \{y \in K : f(x,y) \leq 0\}.$$

Clairement, x^* est une solution du (DEP) si et seulement si $x^* \in \cap_{x\in K} L_f(x)$.

On désignera par K^* et K^d les ensembles-solution du (PEP) et du (DEP) respectivement.

Puisque $K^d = \cap_{x\in K} L_f(x)$, l'ensemble-solution K^d est fermé et convexe si $f(x,\cdot)$ est fermée et convexe sur K.

De même, si $f(x,.)$ est fermée et convexe sur K, alors K^* est fermé convexe.

Si, de plus, $f(\cdot,y)$ est hémicontinue supérieurement, alors $K^d \subseteq K^*$.
En effet, soient x^* solution du (DEP) et un certain $x \in K$.
On définit la combinaison convexe $z_\lambda \in K$ pour tout $\lambda \in]0,1[$ par $z_\lambda = \lambda x + (1-\lambda)x^*$.
Nous avons, puique $f(x,y)$ est une bifonction d'équilibre convexe par rapport à la seconde variable (c'est-à-dire pour chaque $x \in K$ fixé), que pour tout $\lambda \in]0,1[$, $0 = f(z_\lambda, z_\lambda) = f(z_\lambda, \lambda x + [1-\lambda]x^*) \leq \lambda f(z_\lambda, x) + (1-\lambda) f(z_\lambda, x^*)$.
Etant donné que $x^* \in \cap_{x\in K} L_f(x) = \cap_{x\in K}\{y \in K : f(x,y) \leq 0\}$, et comme $z_\lambda \in K$, on obtient, pour tout $\lambda \in]0,1[$, $f(z_\lambda, x^*) \leq 0$.
Ainsi, pour tout $\lambda \in]0,1[$ et pour tout $x \in K$, $0 \leq f(z_\lambda, x) = f(\lambda x + [1-\lambda]x^*, x) = f(x^* + \lambda[x - x^*], x)$.
En passant à la limite supérieure $\lambda \searrow 0$ de part et d'autre de cette inégalité, on conclut, par hémicontinuité supérieure de f par rapport à la première variable, que $0 \leq \lim\ \sup_{\lambda\searrow 0} f(x^* + \lambda[x - x^*], x) = f(x^*, x)$ pour tout $x \in K$, c'est-à-dire que x^* est une solution du (PEP).

Et si, en outre, f est pseudo monotone sur K, alors $K^* = K^d$.
En effet, utilisant la pseudo monotonicité de f, on a $x^* \in K^* \Leftrightarrow x^* \in K$ est tel que $f(x^*, y) \geq 0\ \forall y \in K \Rightarrow x^* \in K$ est tel que $f(y, x^*) \leq 0\ \forall y \in K \Leftrightarrow x^* \in K^d$.

Dans ce qui suit, nous supposerons que $K^d \neq \emptyset$.

Le prochain lemme donne une formulation en point fixe du (PEP).

Lemme 3.2.1 *Soit $f : K \times K \to I\!R \cup \{\infty\}$ une bifonction d'équilibre. Alors, les assertions qui suivent sont équivalentes.*

1. *x^* est une solution du (PEP).*
2. *x^* est une solution du problème*

$$min_{y\in K}\ f(x^*, y).$$

Preuve. *$x^* \in K$ est une solution du problème de $min_{y\in K}\ f(x^*, y)$*

$\Leftrightarrow f(x^*, y) \geq f(x^*, x^*) = 0\ \forall y \in K$

$\Leftrightarrow \exists x^* \in K$ tel que $f(x^*, y) \geq 0\ \forall y \in K$.

Le principal inconvénient de la formulation en point fixe donnée par ce lemme est que le problème de minimisation, en général, peut ne pas avoir de solution, et, s'il en a une, celle-ci peut ne pas être unique. Pour éviter cette situation, il est préférable d'utiliser un autre problème d'équilibre auxiliaire qui est équivalent au (PEP).

Soit $L : K \times K \to \mathbb{R}$ une fonction positive telle que $f(x, .)$ soit différentiable et convexe sur K, et telle que :

1. $L(x, x) = 0$ pour tout $x \in K$,
2. $\nabla_2 L(x, x) = 0$ pour tout $x \in K$.

où, comme d'usage, $\nabla_2 L(x, x)$ représente le gradient de la fonction $L(x, .)$ en x. Un exemple important d'une telle fonction est $L(x, y) := \frac{1}{2}\|y - x\|^2$.

Considérons le problème d'équilibre auxiliaire, noté (AuPEP),

$$\text{trouver } x^* \in K \text{ tel que } \rho f(x^*, y) + L(x^*, y) \geq 0 \text{ pour tout } y \in K$$

où $\rho > 0$ est un paramètre de régularisation. L'équivalence entre (PEP) et (AuPEP) est établie dans le prochain lemme.

Lemme 3.2.2 *Soient $f : K \times K \to \cup\{\infty\}$ une bifonction d'équilibre et $x^* \in K$. Supposons que $f(x^*, .) : K \to \mathbb{R}$ est convexe et admet un sous-différentiel. Soit $L : K \times K \to \mathbb{R}^+$ une fonction telle que $L(x, .)$ soit différentiable et convexe sur K, et telle que*

1. $L(x, x) = 0$ pour tout $x \in K$,

2. $\nabla_2 L(x, x) = 0$ pour tout $x \in K$.

Alors $x^* \in K$ est une solution du (PEP) si et seulement si x^* est une solution du (AuPEP).

Preuve. Evidemment, si $x^* \in K$ est une solution du (PEP), alors, par positivité de la fonction L, x^* est une solution du (AuPEP). Inversément, supposons que $x^* \in K$ est une solution du (AuPEP). Appliquant le dernier lemme à la fonction d'équilibre $\rho f + L$, on voit que x^* est un minimiseur du programme convexe

$$\min_{y \in K} \, [\rho f(x^*, y) + L(x^*, y)].$$

Par la condition nécessaire et suffisante d'optimalité bien connue en programmation convexe, x^* est une solution optimale du dernier problème si et seulement si

$$0 \in \partial_2[\rho f(x^*, x^*) + L(x^*, x^*)] + N_K(x^*),$$

où ∂_2 est le sous-différentiel par rapport au second argument, et où

$$N_K(x^*) := \{w : \langle w, x - x^* \rangle \leq 0, \, \forall x \in K\}$$

est le cône normal (extérieur) de K en $x^* \in K$. Vu que $f(x^*, .)$ admet un sous-différentiel et que L est différentiable, par le théorème de Moreau-Rockafellar (voir [7]),

$$\partial_2[\rho f(x^*, x^*) + L(x^*, x^*)] = \rho \partial_2 f(x^*, x^*) + \nabla_2 L(x^*, x^*).$$

Puisque $\nabla_2 L(x^*, x^*) = 0$ par hypothèse, il s'ensuit que

$$0 \in \partial_2 f(x^*, x^*) + N_K(x^*),$$

ce qui constitue une condition nécessaire et suffisante pour que $x^* \in K$ soit un minimiseur du problème

$$\min_{y \in K} \; f(x^*, y).$$

Appliquant de nouveau le dernier lemme, on conclut que $f(x^*, y) \geq 0$ pour tout $y \in K$.

3.3 L'algorithme de l'extragradient.

Comme nous l'avons mentionné, si $f(x, .)$ est fermée et convexe sur K, et si $f(\cdot, y)$ est hémicontinue supérieurement sur K, alors l'ensemble-solution du (DEP) est contenu dans celui du (PEP). Dans l'algorithme qui va suivre, nous utiliserons la bifonction auxiliaire donnée par

$$L(x, y) := G(y) - G(x) - \langle \nabla G(x), y - x \rangle,$$

où $G : \mathbb{R}^n \to \mathbb{R}$ est une fonction fortement convexe (de module $\beta > 0$) et continuement différentiable, telle que

1. $L(x, x) = 0$ pour tout $x \in K$,
2. $\nabla_2 L(x, x) = 0$ pour tout $x \in K$;

par exemple $G(x) := \frac{1}{2}\|x\|^2$. Dans ce cas, le problème auxiliaire (AuPEP) prend la forme

trouver $x^* \in K$ tel que

$$\rho f(x^*, y) + G(y) - G(x^*) - \langle \nabla G(x^*), y - x^* \rangle \geq \text{ pour tout } y \in K.$$

Puisque $L(x, .)$ est fortement convexe (car somme de fonctions convexe, fortement convexe et linéaire) sur l'ensemble K fermé et convexe, le problème, noté (Cx),

$$\min_{y \in K} \; [\rho f(x, y) + G(y) - G(x) - \langle \nabla G(x), y - x \rangle]$$

admet toujours une unique solution. Le dernier lemme donne une formulation en point fixe pour le problème (PEP) qui suggère une méthode itérative pour résoudre le (PEP) en posant $x^{k+1} := s(x^k)$ où $s(x^k)$ est l'unique solution du problème fortement convexe (Cx^k). Malheureusement, il est bien connu que, pour des problèmes d'inéquations variationnelles monotones qui sont des cas particuliers de problèmes d'équilibre (PEP) monotones, la suite $\{x^k\}$ peut ne pas être convergente. Ce fait est à l'origine l'utilisation de la méthode de l'extragradient pour les inéquations variationnelles monotones. Pour le problème d'inéquation variationnelle univaluée, noté (VIP), donné par

$$\text{trouver } x^* \in K \text{ tel que } \langle F(x^*), x - x^* \rangle \geq 0 \text{ pour tout } x \in K,$$

la méthode de l'extragradient (ou de la double projection) construit deux suites $\{x^k\}$ et $\{y^k\}$ en posant

$$y^k := \Pi_K[x^k - \rho F(x^k)] \text{ et } x^{k+1} := \Pi_K[x^k - \rho F(y^k)]$$

où $\rho > 0$ et Π_K représente la projection euclidienne sur K.

Maintenant, nous allons étendre la méthode de l'extragradient au problème d'équilibre (PEP). Durant le reste de ce chapitre, nous supposerons que, pour tout $x \in K$, la fonction $f(x, .)$ est fermée et convexe, et admet un sous-différentiel sur K. Sous ces hypothèses, les sous-problèmes que l'on devra résoudre dans les prochains algorithmes sont des programmes convexes avec des fonctions objectif fortement convexes. Pour obtenir la convergence du premier algorithme que nous allons décrire, le paramètre de régularisation ρ devra satisfaire une certaine condition (voir le théorème de convergence qui suivra).

Algorithme 3.3.1

Pas 0 : *Prendre $x^0 \in K$, $\rho > 0$ et poser $k := 0$.*

Pas 1 : *Résoudre le programme fortement convexe*

$$min_{y \in K} \; [\rho f(x^k, y) + G(y) - \langle \nabla G(x^k), y - x^k \rangle]$$

pour obtenir son unique solution optimale y^k.

Si $y^k = x^k$, alors stop : x^k est une solution du (PEP). Sinon, aller au Pas 2.

Pas 2 : *Résoudre le programme fortement convexe*

$$min_{y \in K} \; [\rho f(y^k, y) + G(y) - \langle \nabla G(x^k), y - x^k \rangle]$$

pour obtenir son unique solution x^{k+1}.

Pas 3 : *Poser $k := k + 1$; revenir au Pas 1.*

Le lemme qui suit montre que si cet algorithme se termine, alors, en effet, une solution du (PEP) a été trouvée.

Lemme 3.3.1 *Si l'algorithme se termine à un certain point itéré x^k, alors x^k est une solution du (PEP).*

Preuve. Si $y^k = x^k$, alors, puisque $f(x, x) = 0$ pour tout x dans K, nous avons

$$\rho f(x^k, y^k) + G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle = 0.$$

Vu que $y^k = x^k$ est une solution du premier programme fortement convexe qui est inchangé à une constante près, nous avons

$= \rho f(x^k, y^k) + G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle$

$\leq \rho f(x^k, y) + G(y) - G(x^k) - \langle \nabla G(x^k), y - x^k \rangle \ \forall y \in K.$

Donc, par l'équivalence entre (AuPEP) et (PEP), $x^k = y^k$ est une solution du (PEP).

Le prochain théorème établit la convergence de l'algorithme.

Théorème 3.3.1 *Supposons que*

1. *G est fortement convexe de module $\beta > 0$ et continuement différentiable sur un ensemble ouvert Ω contenant K.*
2. *il existe deux constantes $c_1 > 0$ et $c_2 > 0$, telles que*

$$f(x, y) + f(y, z) \geq f(x, z) - c_1\|y - x\|^2 - c_2\|z - y\|^2 \ \forall x, y, z \in K.$$

Alors

1. *pour tout $x^* \in K^d$, on a*

$$l(x^k) - l(x^{k+1}) \geq (\frac{\beta}{2} - \rho c_1)\|y^k - x^k\|^2 + (\frac{\beta}{2} - \rho c_2)\|x^{k+1} - y^k\|^2$$

où $l(y) := G(x^*) - G(y) - \langle \nabla G(y), x^* - y \rangle$ pour tout $y \in K$.

2. *si, de plus, f est semi-continue inférieurement sur $K \times K$, $f(\cdot, y)$ est semi-continue supérieurement, et $0 < \rho < min\ \{\frac{\beta}{2c_1}, \frac{\beta}{2c_2}\}$, la suite $\{x^k\}$ est bornée et chacun de ses points d'adhérence est une solution du (PEP).*
En outre, si $K^d = K^*$ (en particulier, si f est pseudo monotone sur K), la suite $\{x^k\}$ toute entière converge vers une solution du (PEP).

Preuve.

1. Prenons n'importe quel $x^* \in K^d$. Par définition de la fonction l et puisque $x^k, x^{k+1} \in K$, nous avons que

$l(x^k) - l(x^{k+1})$

$= G(x^{k+1}) - G(x^k) + \langle \nabla G(x^{k+1}), x^* - x^{k+1} \rangle - \langle \nabla G(x^k), x^* - x^k \rangle$

$= G(x^{k+1}) - G(x^k) + \langle \nabla G(x^{k+1}) - \nabla G(x^k), x^* - x^{k+1} \rangle - \langle \nabla G(x^k), x^{k+1} - x^k \rangle$ en utilisant un artifice de calcul.

Utilisant la condition nécessaire et suffisante d'optimalité bien connue en programmation convexe, on voit que x^{k+1} résout le programme convexe du Pas 2

$$\min_{y \in K}\ [\rho f(y^k, y) + G(y) - G(x^k) - \langle \nabla G(x^k), y - x^k \rangle],$$

si et seulement si

$$0 \in \partial_2[\rho f(y^k, x^{k+1}) + G(x^{k+1}) - G(x^k) - \langle \nabla G(x^k), x^{k+1} - x^k \rangle] + N_K(x^{k+1})$$

où $N_K(x)$ est le cône normal (extérieur) de K en $x \in K$.Vu que $f(y^k, .)$ admet un sous-différentiel et que G est différentiable, par le théorème de Moreau-Rockafellar (voir [7]), on a que

$$\partial_2[\rho f(y^k, x^{k+1}) + G(x^{k+1}) - G(x^k) - \langle \nabla G(x^k), x^{k+1} - x^k \rangle]$$
$$= \rho \partial_2 f(y^k, x^{k+1}) + \nabla G(x^{k+1}) - \nabla G(x^k).$$

Ainsi, il existe $w \in \partial_2 f(y^k, x^{k+1})$ tel que

$$\langle -\rho w - \nabla G(x^{k+1}) + \nabla G(x^k), y - x^{k+1} \rangle \leq 0 \ \forall y \in K$$
$$\Leftrightarrow \langle \nabla G(x^{k+1}) - \nabla G(x^k), y - x^{k+1} \rangle \geq \rho \langle w, x^{k+1} - y \rangle \ \forall y \in K.$$

Par la définition du sous-gradient w, nous avons, utilisant la dernière inéquation, que

$$\langle \nabla G(x^{k+1}) - \nabla G(x^k), y - x^{k+1} \rangle \geq \rho f(y^k, x^{k+1}) - \rho f(y^k, y) \ \forall y \in K.$$

Avec $y = x^*$, cette inéquation devient

$$\langle \nabla G(x^{k+1}) - \nabla G(x^k), x^* - x^{k+1} \rangle \geq \rho f(y^k, x^{k+1}) - \rho f(y^k, x^*).$$

Puisque x^* est une solution du (DEP), $f(y^k, x^*) \leq 0$. Donc

$$\langle \nabla G(x^{k+1}) - \nabla G(x^k), x^* - x^{k+1} \rangle \geq \rho f(y^k, x^{k+1}).$$

Maintenant, appliquant la seconde hypothèse avec $x = x^k$, $y = y^k$ et $z = x^{k+1}$, il suit de la dernière inéquation que

$$\langle \nabla G(x^{k+1}) - \nabla G(x^k), x^* - x^{k+1} \rangle$$
$$\geq \rho f(x^k, x^{k+1}) - \rho f(x^k, y^k) - \rho c_1 \|y^k - x^k\|^2 - \rho c_2 \|x^{k+1} - y^k\|^2.$$

D'autre part, par le Pas 1, comme y^k est la solution du programme convexe (inchangé à une constante près)

$$\min_{y \in K} \ [\rho f(x^k, y) + G(y) - G(x^k) - \langle \nabla G(x^k), y - x^k \rangle],$$

nous avons, par la condition nécessaire et suffisante d'optimalité bien connue en programmation convexe,

$$0 \in \partial_2[\rho f(x^k, y^k) + G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle] + N_K(y^k).$$

Vu que $f(x^k, .)$ admet un sous-différentiel et que G est différentiable, par le théorème de Moreau-Rockafellar (voir [7]), on a que

$$\partial_2[\rho f(x^k, y^k) + G(y^k) + G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle]$$
$$= \rho \partial_2 f(x^k, y^k) + \nabla G(y^k) - \nabla G(x^k).$$

Ainsi, il existe $w' \in \partial_2 f(x^k, y^k)$ tel que

$$\langle -\rho w' - \nabla G(y^k) + \nabla G(x^k), y - y^k \rangle \leq 0 \ \forall y \in K$$

$$\Leftrightarrow \langle \nabla G(y^k) - \nabla G(x^k), y - y^k\rangle \geq \rho\langle w', y^k - y\rangle \; \forall y \in K.$$

Par la définition du sous-gradient w', nous avons, utilisant la dernière inéquation, que

$$\langle \nabla G(y^k) - \nabla G(x^k), y - y^k\rangle \geq \rho f(x^k, y^k) - \rho f(x^k, y) \; \forall y \in K$$

$$\Leftrightarrow \rho f(x^k, y) - \rho f(x^k, y^k) \geq \langle \nabla G(y^k) - \nabla G(x^k), y^k - y\rangle \; \forall y \in K.$$

Avec $y = x^{k+1}$, on obtient

$$\rho f(x^k, x^{k+1}) - \rho f(x^k, y^k) \geq \langle \nabla G(y^k) - \nabla G(x^k), y^k - x^{k+1}\rangle.$$

Des inégalités qui précèdent, il suit que

$l(x^k) - l(x^{k+1})$

$\geq G(x^{k+1}) - G(x^k) + \rho f(x^k, x^{k+1}) - \rho f(x^k, y^k) - \rho c_1\|y^k - x^k\|^2 - \rho c_2\|x^{k+1} - y^k\|^2 - \langle \nabla G(x^k), x^{k+1} - x^k\rangle$

$\geq G(x^{k+1}) - G(x^k) + \langle \nabla G(y^k) - \nabla G(x^k), y^k - x^{k+1}\rangle - \rho c_1\|y^k - x^k\|^2 - \rho c_2\|x^{k+1} - y^k\|^2 - \langle \nabla G(x^k), x^{k+1} - x^k\rangle$

$= G(x^{k+1}) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle + \langle \nabla G(y^k), y^k - x^{k+1}\rangle - \rho c_1\|y^k - x^k\|^2 - \rho c_2\|x^{k+1} - y^k\|^2$

$= [G(x^{k+1}) - G(y^k) - \langle \nabla G(y^k), x^{k+1} - y^k\rangle] + [G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle] - \rho c_1\|y^k - x^k\|^2 - \rho c_2\|x^{k+1} - y^k\|^2$ en utilisant un artifice de calcul.

Puisque G est fortement convexe de module $\beta > 0$, pour tout x et y, on a, par définition,

$$G(y) - G(x) - \langle \nabla G(x), y - x\rangle \geq \frac{\beta}{2}\|y - x\|^2 \; \forall x, y \in K.$$

Appliquant cette propriété tout d'abord à y^k et x^{k+1}, et ensuite à x^k et y^k, on obtient que

$$l(x^k) - l(x^{k+1}) \geq (\frac{\beta}{2} - \rho c_1)\|y^k - x^k\|^2 + (\frac{\beta}{2} - \rho c_2)\|x^{k+1} - y^k\|^2 \; \forall k,$$

ce qui prouve la première assertion.

2. Par l'hypothèse $0 < \rho < min\{\frac{\beta}{2c_1}, \frac{\beta}{2c_2}\}$, nous avons

$$\frac{\beta}{2} - \rho c_1 > 0 \text{ et } \frac{\beta}{2} - \rho c_2 > 0.$$

Donc, utilisant la première assertion, on déduit que

$$l(x^k) - l(x^{k+1}) \geq (\frac{\beta}{2} - \rho c_1)\|y^k - x^k\|^2 \geq 0 \; \forall k.$$

Dès lors, la suite $\{l(x^k)\}_{k\geq 0}$ est décroissante. Puisqu'elle est bornée inférieurement par 0, elle est convergente. Utilisant la dernière inégalité pour tout k de 0 à m, et passant à la somme, nous avons

$$(\frac{\beta}{2} - \rho c_1)\sum_{k=0}^{m} \|y^k - x^k\|^2 \leq l(x^0) - l(x^{m+1}) \ \forall m \geq 0.$$

Comme la suite $\{l(x^k)\}_{k\geq 0}$ est convergente, en passant à la limite $m \to \infty$, nous obtenons que

$$\sum_{k=0}^{\infty} \|y^k - x^k\|^2 < \infty$$

ce qui implique

$$\lim_{k\to\infty} \|y^k - x^k\| = 0$$

et, par unicité de la limite, que

$$\lim_{k(\in\kappa)\to\infty} \|y^k - x^k\| = 0$$

pour toute sous-suite $\{x^k : k \in \kappa\}$.

Notons que, par convexité forte de G de module β et définition de $l(x^k)$, nous pouvons écrire

$$0 \leq \frac{\beta}{2}\|x^* - x^k\|^2 \leq l(x^k) \ \forall k.$$

Dès lors, puisque la suite $\{l(x^k)\}_{k\geq}$ est convergente et donc bornée, nous pouvons déduire que la suite $\{x^k\}_{k\geq 0}$ est bornée, de sorte que, par le théorème de Bolzano-Weierstrass, elle admet au moins un point d'adhérence. Soit $\bar{x} \in K$ un point d'adhérence de $\{x^k\}_{k\geq 0}$ et $\{x^{k_i}\}_{i\geq 0}$ la sous-suite telle que

$$\lim_{i\to\infty} x^{k_i} = \bar{x}.$$

Donc, puisque $\lim_{i\to\infty} \|y^{k_i} - x^{k_i}\| = 0$, il suit que

$$\lim_{i\to\infty} y^{k_i} = \bar{x}.$$

De nouveau, par le Pas 1 de l'algorithme, et puisque la minimisation se fait à une constante près, nous avons

$$\rho f(x^{k_i}, y) + G(y) - G(x^{k_i}) - \langle \nabla G(x^{k_i}), y - x^{k_i}\rangle$$

$$\geq \rho f(x^{k_i}, y^{k_i}) + G(y^{k_i}) - G(x^{k_i}) - \langle \nabla G(x^{k_i}), y^{k_i} - x^{k_i}\rangle \ \forall y \in K.$$

Etant donné que $f(\cdot, y)$ est semi-continue supérieurement, que f est semi-continue inférieurement et que $f(\bar{x}, \bar{x}) = 0$, en passant à la limite $i \to \infty$, nous obtenons, à partir de la dernière inéquation, que $\forall y \in K$

$$\rho f(\bar{x}, y) + G(y) - G(\bar{x}) - \langle \nabla G(\bar{x}), y - \bar{x}\rangle$$

$$\geq \limsup_{i\to\infty} [\rho f(x^{k_i}, y) + G(y) - G(x^{k_i}) - \langle \nabla G(x^{k_i}), y - x^{k_i}\rangle]$$

$$\geq \limsup_{i\to\infty} [\rho f(x^{k_i}, y^{k_i}) + G(y^{k_i}) - G(x^{k_i}) - \langle \nabla G(x^{k_i}), y^{k_i} - x^{k_i}\rangle]$$

$$\geq \liminf_{i\to\infty} [\rho f(x^{k_i}, y^{k_i}) + G(y^{k_i}) - G(x^{k_i}) - \langle \nabla G(x^{k_i}), y^{k_i} - x^{k_i}\rangle]$$

$$\geq \rho f(\bar{x}, \bar{x}) + G(\bar{x}) - G(\bar{x}) - \langle \nabla G(\bar{x}), \bar{x} - \bar{x}\rangle$$

$$= 0,$$

ce qui montre que $\bar{x}$ est une solution du (AuPEP) correspondant à $L(x, y) = G(y) - G(x) - \langle \nabla G(x), y - x\rangle$. Donc, par l'équivalence entre (AuPEP) et (PEP), $\bar{x}$ est une solution du (PEP).
En supposant maintenant que $K^d = K^*$, montrons que la suite $\{x^k\}_{k\geq 0}$ toute entière converge vers $\bar{x}$. Utilisant la définition de $l(y)$ avec $x^* = \bar{x} \in K^d$, nous avons $l(\bar{x}) = 0$. Donc, puisque G est fortement convexe de module β, nous pouvons écrire

$$l(x^k) - l(\bar{x}) = G(\bar{x}) - G(x^k) - \langle \nabla G(x^k), \bar{x} - x^k\rangle \geq \frac{\beta}{2}\|x^k - \bar{x}\|^2 \geq 0 \ \forall k \geq 0.$$

D'autre part, étant donné que la suite $\{l(x^k)\}_{k\geq 0}$ est positive et décroissante, donc convergente, et que, par différentiabilité continue de G, $l(x^{k_i}) \to l(\bar{x})$, nous devons avoir, par unicité de la limite, $l(x^k) \to l(\bar{x})$ quand $k \to \infty$. Ainsi, appliquant le théorème de l'étau aux dernières inégalités, $\lim_{k\to\infty} x^k = \bar{x} \in K^*$.

Remarque 3.3.1 *La seconde hypothèse du précédent théorème n'implique pas nécessairement que f est continue. En fait, si $f(x, y) := \varphi(y) - \varphi(x)$, cette hypothèse devient*

$$\varphi(y) - \varphi(x) + \varphi(z) - \varphi(y) \geq \varphi(z) - \varphi(x) - c_1\|y - x\|^2 - c_2\|z - y\|^2$$

$$\Leftrightarrow 0 \geq -c_1\|y - x\|^2 - c_2\|z - y\|^2$$

et tient pour tous $c_1, c_2 \geq 0$ et pour toute fonction φ.

3.4 Les algorithmes de recherche linéaire.

L'algorithme que nous venons de voir requiert que f satisfasse une condition de type Lipschitz qui, parfois, est inconnue. Afin d'éviter cette exigence, dans cette section, nous allons modifier ce dernier par l'utilisation d'une recherche linéaire. La technique de recherche linéaire a été largement utilisée dans des méthodes de descente, aussi bien pour des problèmes de programmation mathématique que pour des inéquations variationnelles.

Commençons tout d'abord par la définition qui suit.

Définition 3.4.1 *Soit K un ensemble non vide et fermé dans $I\!R^n$. Un opérateur $P : I\!R^n \to I\!R^n$ est dit*

1. *réalisable par rapport à K si et seulement si*

$$P(x) \in K \ \forall x \in I\!R^n,$$

2. *quasi non expansif par rapport à K si et seulement si, $\forall x \in I\!R^n$,*

$$\|P(x) - y\| \leq \|x - y\| \ \forall y \in K.$$

Notons que, si $\Pi_K(\cdot)$ est la projection euclidienne sur K, alors $\Pi_K(\cdot)$ est un opérateur réalisable quasi non expansif par rapport à K. On représente par $\mathcal{F}$(K) la classe des opérateurs réalisables quasi non expansifs par rapport à K.

A partir de maintenant, nous choisirons une suite $\{\gamma_k\}_{k\geq 0}$ telle que

$$\gamma_k \in]0,2[\ \forall k = 0,1,... \quad \text{et} \quad \liminf \gamma_k(2-\gamma_k) > 0.$$

L'algorithme peut alors être décrit comme suit.

Algorithme 3.4.1

Données : *$x^0 \in K$, $\alpha \in]0,1[$, $\theta \in]0,1[$ et $\rho > 0$.*

Pas 0 : *Poser $k := 0$.*

Pas 1 : *Résoudre le programme fortement convexe*

$$min_{y\in K} \ \left(f(x^k, y) + \frac{1}{\rho}[G(y) - \langle \nabla G(x^k), y - x^k \rangle] \right)$$

pour obtenir son unique solution y^k.

Si $y^k = x^k$, alors stop : x^k est une solution du (PEP). Sinon, aller au Pas 2.

Pas 2 :

Pas 2.1 : trouver le plus petit entier strictement positif m tel que

$$\begin{cases} z^{k,m} = (1-\theta^m)x^k + \theta^m y^k, \\ f(z^{k,m}, y^k) + \frac{\alpha}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle] \leq 0. \end{cases}$$

Pas 2.2 : poser $\theta_k := \theta^m$ et $z^k := z^{k,m}$ Si $0 \in \partial_2 f(z^k, z^k)$, alors stop : z^k est une solution du (PEP). Sinon, aller au Pas 3.

Pas 3 : *Prendre $g^k \in \partial_2 f(z^k, z^k)$,*

$$\sigma_k = \frac{-\theta_k f(z^k, y^k)}{(1-\theta_k)\|g^k\|^2} \quad \text{et} \quad x^{k+1} = P_k(x^k - \gamma_k \sigma_k g^k)$$

où $P_k \in \mathcal{F}(K)$.

Pas 4 : *Poser $k := k+1$; revenir au Pas 1.*

Le prochain lemme stipule que si cet algorithme termine aux Pas 1 ou 2.2, alors, en effet, une solution du (PEP) a été trouvée.

Lemme 3.4.1 *Si l'algorithme se termine au Pas 1 (respectivement 2.2), alors x^k (respectivement z^k) est une solution du (PEP).*

Preuve. Si l'algorithme se termine au Pas 1, alors $x^k = y^k$. Puisque y^k est la solution du programme fortement convexe qui est inchangé à une constante et une constante multiplicative près, nous avons, $\forall y \in K$,

$$\rho f(x^k, y) + G(y) - G(x^k) - \langle \nabla G(x^k), y - x^k \rangle$$

$$\rho f(x^k, x^k) + G(x^k) - G(x^k) - \langle \nabla G(x^k), x^k - x^k \rangle$$

$$= 0.$$

Donc, par l'équivalence entre (AuPEP) et (PEP), x^k est une solution du (PEP). Si l'algorithme se termine au Pas 2.2, alors $0 \in \partial_2 f(z^k, z^k)$. Puisque $f(z^k, .)$ est en outre convexe, on a $f(z^k, z^k) \leq f(z^k, y)$ pour tout $y \in K$. Ainsi, le fait que $f(z^k, z^k) = 0$ montre que z^k est une solution du (PEP).

Le lemme qui suit montre qu'il existe toujours un entier strictement positif m tel que la condition du Pas 2.1 est satisfaite.

Lemme 3.4.2 *Supposons que $f(\cdot, y)$ est semi-continue supérieurement et que $y^k \neq x^k$. Alors*

1. *il existe un entier fini $m > 0$ tel que l'inégalité du Pas 2.1 tient.*
2. *$f(z^k, y^k) < 0$.*

Preuve.

1. Supposons par l'absurde que pour tout entier strictement positif m et

$$z^{k,m} = (1 - \theta^m)x^k + \theta^m y^k,$$

nous avons

$$f(z^{k,m}, y^k) + \frac{\alpha}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle] > 0.$$

En passant à la limite $m \to \infty$, puisque $f(\cdot, y^k)$ est semi-continue supérieurement, nous avons

$$f(x^k, y^k) + \tfrac{\alpha}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle]$$

$$\geq \limsup_{m \to \infty} \left(f(z^{k,m}, y^k) + \tfrac{\alpha}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle] \right)$$

$$\geq 0.$$

D'autre part, puisque y^k est une solution du programme fortement convexe du Pas 1 qui est inchangé à une constante près, nous pouvons écrire, $\forall y \in K$,

$f(x^k, y^k) + \frac{1}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle]$

$\leq f(x^k, y) + \frac{1}{\rho}[G(y) - G(x^k) - \langle \nabla G(x^k), y - x^k \rangle]$.

Avec $y = x^k$, la dernière inéquation devient, puisque f est une bifonction d'équilibre,

$$f(x^k, y^k) + \frac{1}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle] \leq 0.$$

Ainsi, utilisant les deux dernières inégalités,

$f(x^k, y^k) + \frac{1}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle]$

$\leq f(x^k, y^k) + \frac{\alpha}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle]$

Vu qu'en outre, par la convexité forte de G,

$$G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle \geq 0,$$

nous déduisons que soit $G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle = 0$, soit $\alpha \geq 1$. Le premier cas implique que $x^k = y^k$, puisque G est fortement convexe. D'où, puisque par hypothèse $\alpha \in]0, 1[$, les deux cas mènent à une contradiction.

2. Au vu de sa règle de détermination, z^k vérifie

$$f(z^k, y^k) + \frac{\alpha}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle] \leq 0.$$

Or, par la convexité forte de module $\beta > 0$ et par le fait que $x^k \neq y^k$,

$$G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle \geq \frac{\beta}{2}\|x^k - y^k\|^2 > 0.$$

Ainsi,

$$f(z^k, y^k) \leq -\frac{\alpha}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle] \leq -\frac{\alpha\beta}{2\rho}\|x^k - y^k\|^2 < 0.$$

Afin de prouver la convergence de l'algorithme, nous donnons la propriété clé qui suit de la suite $\{x^k\}_{k \geq 0}$ générée par l'algorithme.

Lemme 3.4.3 *Si $f(x, \cdot)$ est convexe et admet un sous-différentiel, alors les assertions suivantes tiennent.*

1. *Pour toute solution x^* du (DEP), on a*

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2 - \gamma_k(2 - \gamma_k)(\sigma_k \|g^k\|)^2.$$

2. $\sum_{k=0}^{\infty} \gamma_k(2 - \gamma_k)(\sigma_k\|g^k\|)^2 < \infty$.
3. *Supposons que l'algorithme ne se termine pas. Alors, si, en outre, $f(x, \cdot)$ est finie sur un ensemble ouvert Ω contenant K et $f(\cdot, y)$ est continue, la suite $\{g^k\}$ est bornée.*

Preuve.

1. Prenons n'importe quel $x^* \in K^d$. Par la définition de x^{k+1} au Pas 3 qui est telle qu'en posant $w^k := x^k - \gamma_k \sigma_k g^k$, $x^{k+1} =: P_k(w^k)$, et par la propriété de quasi non expansivité par rapport à K de P_k, nous avons que

$$\|x^{k+1} - x^*\|^2$$

$$= \|P_k(w^k) - x^*\|^2$$

$$\leq \|w^k - x^*\|^2$$

$$= \|x^k - \gamma_k \sigma_k g^k - x^*\|^2$$

$$= \|x^k - x^*\|^2 - 2\gamma_k \sigma_k \langle g^k, x^k - x^* \rangle + (\gamma_k \sigma_k \|g^k\|)^2.$$

Etant donné que $g^k \in \partial_2 f(z^k, z^k)$, et que $f(z^k, .)$ est convexe, nous avons

$$\langle g^k, x^k - x^* \rangle$$

$= \langle g^k, x^k - z^k \rangle + \langle g^k, z^k - x^* \rangle$ en utilisant un artifice de calcul

$$\geq \langle g^k, x^k - z^k \rangle + f(z^k, z^k) - f(z^k, x^*)$$

$= \langle g^k, x^k - z^k \rangle - f(z^k, x^*)$ puisque f est une bifonction d'équilibre

$\geq \langle g^k, x^k - z^k \rangle$ puisque $x^* \in K^d$.

Utilisant le Pas 2, nous pouvons écrire

$$(1 - \theta_k) x^k - z^k = -\theta_k y^k$$

$$\Leftrightarrow x^k - \tfrac{1}{1-\theta_k} z^k = -\tfrac{\theta_k}{1-\theta_k} y^k$$

$\Leftrightarrow x^k - \frac{1-\theta_k}{1-\theta_k} z^k = \frac{\theta_k}{1-\theta_k} z^k - \frac{\theta_k}{1-\theta_k} y^k$ en utilisant un artifice de calcul

$$\Leftrightarrow x^k - z^k = \tfrac{\theta_k}{1-\theta_k}(z^k - y^k),$$

de sorte que

$$\langle g^k, x^k - z^k \rangle$$

$$= \tfrac{\theta_k}{1-\theta_k} \langle g^k, z^k - y^k \rangle$$

$\geq \frac{\theta_k}{1-\theta_k} [f(z^k, z^k) - f(z^k, y^k)]$ étant donné que $g^k \in \partial_2 f(z^k, z^k)$ et que $f(z^k,)$ est convexe

$= \frac{-\theta_k}{1-\theta_k} f(z^k, y^k)$ puisque f est une bifonction d'équilibre.

Par la seconde assertion du précédent lemme, le Pas 3 et le fait que $\theta_k \in]0,1[$, il suit que

$$\frac{-\theta_k}{1-\theta_k} f(z^k, y^k) = \sigma_k \|g^k\|^2 > 0$$

qui, en utilisant, les relations démontrées plus haut, implique, $\forall x^* \in K^d$,

$\|x^{k+1} - x^*\|^2$

$\leq \|x^k - x^*\|^2 - 2\gamma_k \sigma_k \langle g^k, x^k - x^* \rangle + (\gamma_k \sigma_k \|g^k\|)^2$

$\leq \|x^k - x^*\|^2 - 2\gamma_k \sigma_k^2 \|g^k\|^2 + (\gamma_k \sigma_k \|g^k\|)^2$

$= \|x^k - x^*\|^2 - \gamma_k(2-\gamma_k)(\sigma_k \|g^k\|)^2.$

2. En appliquant la dernière inéquation pour tout k de 0 à m, nous obtenons

$$\sum_{k=0}^{m} \gamma_k(2-\gamma_k)(\sigma_k \|g^k\|)^2 \leq \|x^0 - x^*\|^2 - \|x^{m+1} - x^*\|^2.$$

Puisque la suite positive $\{\|x^m - x^*\|\}_{m \geq 0}$ est décroissante par la première assertion et la première hypothèse sur γ_k, et donc convergente, en passant à la limite $m \to \infty$, nous obtenons

$$\sum_{k=0}^{\infty} \gamma_k(2-\gamma_k)(\sigma_k \|g^k\|)^2 < \infty$$

3. Observons tout d'abord que $\{y^k\}$ est bornée. En effet, puisque y^k est l'unique solution du programme fortement convexe du Pas 1 dont la fonction objectif est continue et l'ensemble réalisable est constant, par le théorème maximal (voir Proposition 23 dans [1]), l'opérateur $s(x^k) : x^k \hookrightarrow y^k$ est continu. Puisque $\{x^k\}$ est bornée par la première assertion, $\{y^k\}$ est bornée, et, par conséquent, $\{z^k\}$ est également bornée, puisque z^k est combinaison convexe de x^k et y^k. Donc, par passage au théorème de Bolzano-Weierstrass et donc à une sous suite de $\{z^k\}$ et à son point d'adhérence z^*, $\lim_{k\to\infty} z^k = z^*$ où la sous-suite est, sans perte de généralité, indexée comme la suite. Par continuité par rapport à la première variable de la fonction convexe $f(z^k, .)$, la suite $\{f(z^k, .)\}$ converge simplement vers $f(z^*, .)$. Etant donné qu'en outre, f est convexe et finie par rapport à la seconde variable, et que $g^k \in \partial_2 f(z^k, z^k)$ par le Pas 3, nous pouvons déduire que $\{g^k\}$ est bornée

Nous sommes maintenant en mesure de prouver le prochain théorème de convergence de l'algorithme. Comme nous l'avons vu dans le premier lemme, si l'algorithme se termine, alors une solution du (PEP) a été trouvée. Sinon, nous avons les résultats de convergence qui suivent.

Théorème 3.4.1 *En plus des hypothèses des deux derniers lemmes, nous supposons que $f(x,.)$ est continue. Alors*

1. *la suite $\{x^k\}$ est bornée et chacun de ses points d'adhérence est une solution du (PEP).*
2. *si $K^d = K^*$ (en particulier, si f est pseudo monotone), alors la suite $\{x^k\}$ toute entière converge vers une solution du (PEP). En outre, si $\gamma_k = \gamma \in]0,2[$ pour tout $k \geq 0$, alors*

$$lim\ inf_{k\to\infty}\ (\sigma_k\|g^k\|\sqrt{k+1}) = 0.$$

Preuve.

1. Le caractère borné de $\{x^k\}$ suit immédiatement de la première assertion du précédent lemme. Par la seconde assertion de ce même lemme, nous avons

$$\gamma_k(2-\gamma_k)(\sigma_k\|g^k\|)^2 \to 0 \text{ quand } k \to \infty.$$

Par la seconde hypothèse sur γ_k, $\liminf_{k\to\infty} \gamma_k(2-\gamma_k) > 0$. Donc, $\sigma_k\|g^k\| \to 0$ quand $k \to \infty$. Utilisant le Pas 3,

$$\sigma_k\|g^k\| = \frac{-\theta_k}{(1-\theta_k)\|g^k\|} f(z^k, y^k) \to 0.$$

Puisque $\{g^k\}$ est bornée par la dernière assertion du lemme qui précède, nous déduisons que

$$\frac{-\theta_k}{1-\theta_k} f(z^k, y^k) \to 0 \text{ quand } k \to \infty.$$

D'autre part, puisque G est fortement convexe de module β, d'après le Pas 2.1, nous avons

$$\begin{aligned}
&0\\
&\leq \tfrac{\theta_k}{1-\theta_k}\tfrac{\alpha\beta}{2\rho}\|x^k - y^k\|^2\\
&\leq \tfrac{\theta_k}{1-\theta_k}\tfrac{\alpha}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle]\\
&\leq -\tfrac{\theta_k}{1-\theta_k} f(z^k, y^k).
\end{aligned}$$

Soit alors x^* un point d'adhérence arbitraire de $\{x^k\}$. Supposons que la suite $\{x^k : k \in N \subseteq \mathbb{N}\}$ converge vers x^*.

Nous considérons deux cas :

Cas 1 : $\limsup_{k\in N} \theta_k > 0$. Alors il existe $\bar{\theta} > 0$ et une sous-suite $N^* \subseteq N$ telle que $\theta_k \geq \bar{\theta}$ pour tout $k \in N^*$. En appliquant le théorème de l'étau aux inégalités démontrées plus haut, nous déduisons

$$\lim_{k(\in N^*)\to\infty} \|y^k - x^k\| = 0.$$

Ainsi, la sous-suite $\{y^k : k \in N^*\}$ converge aussi vers x^*. D'où, par le Pas 1, puisque y^k est la solution du programme fortement convexe qui est inchangé à une constante multiplicative près, nous avons, $\forall y \in K$,

$$\rho f(x^k, y^k) + [G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle]$$

$$\leq \rho f(x^k, y) + [G(y) - G(x^k) - \langle \nabla G(x^k), y - x^k \rangle].$$

Passant à la limite $k(\in N^*) \to \infty$, par continuité de f, nous obtenons, $\forall y \in K$,

$$0$$

$$= \rho f(x^*, x^*) + [G(x^*) - G(x^*) - \langle \nabla G(x^*), x^* - x^* \rangle]$$

$$\leq \rho f(x^*, y) + [G(y) - G(x^*) - \langle \nabla G(x^*), y - x^* \rangle].$$

Donc, par l'équivalence entre (AuPEP) et (PEP), x^* est une solution du (PEP).

Cas 2 : $\lim_{k\to\infty} \theta_k = 0$. D'après le Pas 2 de l'algorithme, nous avons

$$z^k = (1-\theta_k)x^k + \theta_k y^k.$$

Puisque y^k est la solution du programme fortement convexe du Pas 1 dont la fonction objectif est continue et l'ensemble réalisable est constant, par le théorème maximal (voir Proposition 23 dans [1]), la suite $\{y^k : k \in N\}$ est bornée.

Donc, par passage au théorème de Bolzano-Weierstrass, nous pouvons affirmer que la sous-suite $\{y^k : k \in N^*\}$ converge vers un certain point d'adhérence y^*. Par la définition de y^k au Pas 1, et puisqu'une minimisation se fait à une constante près, nous avons, pour tout $y \in K$,

$$\rho f(x^k, y^k) + [G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle]$$

$$\leq \rho f(x^k, y) + [G(y) - G(x^k) - \langle \nabla G(x^k), y - x^k \rangle].$$

Passant à la limite $k(\in N^*) \to \infty$, étant donné que, par unicité de la limite, la sous-suite $\{x^k : k \in N^*\}$ converge vers x^*, que f est semi-continue inférieurement, et que $f(\cdot, y)$ est semi-continue supérieurement, nous avons, $\forall y \in K$,

$$\rho f(x^*, y^*) + [G(y^*) - G(x^*) - \langle \nabla G(x^*), y^* - x^* \rangle]$$

$$\leq \liminf_{k(\in N^*)\to\infty} \rho f(x^k, y^k) + [G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle]$$

$$\leq \limsup_{k(\in N^*)\to\infty} \rho f(x^k, y^k) + [G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle]$$

$$\leq \limsup_{k(\in N^*)\to\infty} \rho f(x^k, y) + [G(y) - G(x^k) - \langle \nabla G(x^k), y - x^k \rangle]$$

$\leq \rho f(x^*, y) + [G(y) - G(x^*) - \langle \nabla G(x^*), y - x^* \rangle].$

En substituant $y = x^*$, nous obtenons, puisque f est une bifonction d'équilibre,

$\rho f(x^*, y^*) + [G(y^*) - G(x^*) - \langle \nabla G(x^*), y^* - x^* \rangle]$

$\leq \rho f(x^*, x^*) + [G(x^*) - G(x^*) - \langle \nabla G(x^*), x^* - x^* \rangle].$

$= 0$

D'autre part, étant donné que m est le plus petit entier strictement positif vérifiant la règle du Pas 2.1, nous avons, puisque $\rho > 0$,

$$\rho f(z^{k,m-1}, y^k) + \alpha[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle] > 0.$$

Par unicité de la limite, nous pouvons affirmer que $\theta_k \to 0$ quand $k(\in N^*) \to \infty$, de sorte que $z^{k,m-1} = (1 - \theta^{m-1})x^k + \theta^{m-1}y^k = (1 - \frac{\theta_k}{\theta})x^k + \frac{\theta_k}{\theta}y^k \to x^*$ quand $k(\in N^*) \to \infty$. Par la continuité de f, nous obtenons, en passant à la limite $k(\in N^*) \to \infty$,

$$\rho f(x^*, y^*) + \alpha[G(y^*) - G(x^*) - \langle \nabla G(x^*), y^* - x^* \rangle] \geq 0.$$

Ainsi, nous obtenons

$\rho f(x^*, y^*) + [G(y^*) - G(x^*) - \langle \nabla G(x^*), y^* - x^* \rangle] \leq \rho f(x^*, y^*) + \alpha[G(y^*) - G(x^*) - \langle \nabla G(x^*), y^* - x^* \rangle]$

$\Leftrightarrow (1 - \alpha)[G(y^*) - G(x^*) - \langle \nabla G(x^*), y^* - x^* \rangle] \leq 0.$

Etant donné, de plus, que $G(y^*) - G(x^*) - \langle \nabla G(x^*), y^* - x^* \rangle \geq \frac{\beta}{2}\|y^* - x^*\|^2$ et que $\alpha \in]0, 1[$, nous déduisons

$$(1 - \alpha)\frac{\beta}{2}\|y^* - x^*\|^2 \leq (1 - \alpha)[G(y^*) - G(x^*) - \langle \nabla G(x^*), y^* - x^* \rangle] \leq 0,$$

ce qui, à nouveau puisque $\alpha \in]0, 1[$, implique que $x^* = y^*$. D'où, par l'inéquation démontrée plus haut, $x^* = y^*$ est une solution optimale du programme fortement convexe

$$\min_{y \in K} \left(\rho f(x^*, y) + [G(y) - G(x^*) - \langle \nabla G(x^*), y - x^* \rangle]\right).$$

Donc, de nouveau par l'équivalence entre (AuPEP) et (PEP), x^* est une solution du (PEP).

2. Maintenant, nous supposons que $K^d = K^*$. Comme nous venons de le voir, la suite $\{x^k\}$ a un point d'adhérence $x^* \in K^*$. Puisque $K^d = K^*$, $x^* \in K^d$. Appliquant la première assertion du lemme qui précède, nous voyons que la suite $\{\|x^k - x^*\|\}$ toute entière est convergente.

D'où la suite $\{x^k\}$ doit converger vers x^* puisqu'elle a un sous-suite qui converge vers x^*.
Finalement, supposons par l'absurde qu'avec $\gamma_k = \gamma \in]0,2[$ pour tout $k \geq 0$, $\lim\inf_{k\to\infty} (\sigma_k \|g^k\| \sqrt{k+1}) \neq 0$. Alors, il existe une constante $\mu > 0$ telle que $\sigma_k \|g^k\| \geq \frac{\mu}{\sqrt{k+1}}$ pour tout k. Utilisant la seconde assertion du lemme qui précède et l'hypothèse sur la suite $\{\gamma_k\}_{k\geq 0}$, nous avons

$$\gamma(2-\gamma)\mu^2 \sum_{k=0}^{\infty} \frac{1}{k+1} \leq \gamma(2-\gamma) \sum_{k=0}^{\infty} (\sigma_k \|g^k\|)^2 < \infty,$$

ce qui est impossible. La preuve est ainsi achevée.

Remarque 3.4.1 *En pratique, pour implémenter l'algorithme, on prend une tolérance $\epsilon > 0$ et on termine l'algorithme quand soit $\|x^k - y^k\| \leq \epsilon$, soit $\|g^k\| \leq \epsilon$.*

Remarque 3.4.2 *Si $f(\cdot, y)$ est convexe, alors la règle du Pas 2.1 pour déterminer z^k tient pour tout θ_k satisfaisant $0 < \theta_k \leq 1 - \alpha$ pour tout k.*

En effet, par convexité de $f(\cdot, y)$, nous pouvons écrire

$f(z^k, y^k) + \frac{\alpha}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle]$

$= f([1-\theta_k]x^k + \theta_k y^k, y^k) + \frac{\alpha}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle]$ *par définition de z^k au Pas 2*

$\leq (1-\theta_k) f(x^k, y^k) + \theta_k f(y^k, y^k) + \frac{\alpha}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle]$

$= (1-\theta_k) f(x^k, y^k) + \frac{\alpha}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle]$ *puisque f est une fonction d'équilibre.*

D'autre part, étant donné que y^k est une solution du programme fortement convexe du Pas 1 qui est inchangé à une constante près, nous avons en particulier que

$f(x^k, y^k) + \frac{1}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle]$

$\leq f(x^k, x^k) + \frac{1}{\rho}[G(x^k) - G(x^k) - \langle \nabla G(x^k), x^k - x^k \rangle]$

$= 0$ *puisque f est une fonction d'équilibre.*

Ou, de manière équivalente,

$f(x^k, y^k)$

$\leq \frac{-1}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle]$.

Ainsi,

$f(z^k, y^k) + \frac{\alpha}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle]$

$\leq (1 - \theta_k)\frac{-1}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle] + \frac{\alpha}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle]$

$= \frac{(\alpha - 1 + \theta_k)}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle]$

≤ 0 *par la convexité forte de G et l'hypothèse sur $\{\theta_k\}_{k\geq 0}$;*

ceci montre que la condition de recherche linéaire du Pas 2.1 est satisfaite.

Par cette remarque, lorsque $f(\cdot, y)$ est convexe (par exemple, lorsque f est une fonction selle), afin d'implémenter le Pas 2, nous pouvons simplement prendre $z^k = (1 - \theta_k)x^k + \theta_k y^k$ où θ_k peut être n'importe quelle constante entre zéro et $1 - \alpha$. Dans le cas général, la recherche linéaire du Pas 2.1 mène parfois à ce que θ_k tende vers 0, ce qui peut causer des zigzags. Pour éviter ce phénomène, nous proposons une autre recherche linéaire et obtenons l'algorithme qui suit.

Algorithme 3.4.2

Données : $x^0 \in K$, $\alpha \in]0, 1[$, $\theta \in]0, 1[$ *et* $\rho > 0$.

Pas 0 : *Poser* $k := 0$.

Pas 1 : *Résoudre le programme fortement convexe*

$$min_{y \in K} \left(f(x^k, y) + \frac{1}{\rho}[G(y) - \langle \nabla G(x^k), y - x^k \rangle] \right)$$

pour obtenir son unique solution y^k.

Si $y^k = x^k$, *alors stop :* x^k *est une solution du (PEP). Sinon, aller au Pas 2.*

Pas 2 :

Pas 2.1 : trouver le plus petit entier strictement positif m *tel que*

$$\begin{cases} z^{k,m} = (1 - \theta^m)x^k + \theta^m y^k, \\ f(z^{k,m}, x^k) - f(z^{k,m}, y^k) \geq \frac{\alpha}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle]. \end{cases}$$

Pas 2.2 : poser $\theta_k := \theta^m$ *et* $z^k := z^{k,m}$, *et aller au Pas 3.*

Pas 3 : *Prendre* $g^k \in \partial_2 f(z^k, x^k)$,

$$\sigma_k = \frac{f(z^k, x^k)}{\|g^k\|^2} \quad et \quad x^{k+1} = P_k(x^k - \gamma_k \sigma_k g^k)$$

où $P_k \in \mathcal{F}(K)$.

Pas 4 : *Poser* $k := k+1$; *revenir au Pas 1.*

Comme pour c'était déjà le cas pour l'algorithme qui le précède, pour celui-ci aussi nous avons le lemme qui suit.

Lemme 3.4.4 *Supposons que* $f(\cdot, y)$ *est continue et que* $y^k \neq x^k$. *Alors*

1. *il existe un entier fini* $m > 0$ *tel que l'inégalité du Pas 2.1 tient.*
2. $f(z^k, x^k) > 0$.
3. $0 \notin \partial_2 f(z^k, x^k)$.

Preuve.

1. Supposons par l'absurde que, pour tout entier strictement positif m, nous avons

$$\begin{cases} z^{k,m} = (1-\theta^m)x^k + \theta^m y^k, \\ f(z^{k,m}, x^k) - f(z^{k,m}, y^k) < \frac{\alpha}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle]. \end{cases}$$

En passant à la limite $m \to \infty$, de sorte que $\theta^m \to 0$, et étant donné que $f(\cdot, y^k)$ est continue et que f est une bifonction d'équilibre, nous avons

$$0 = f(x^k, x^k) \leq f(x^k, y^k) + \frac{\alpha}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle].$$

D'autre part, puisque y^k est une solution du programme fortement convexe du Pas 1 qui est inchangé à une constante près, nous pouvons écrire, $\forall y \in K$,

$$f(x^k, y^k) + \tfrac{1}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle]$$

$$\leq f(x^k, y) + \tfrac{1}{\rho}[G(y) - G(x^k) - \langle \nabla G(x^k), y - x^k \rangle].$$

Avec $y = x^k$, la dernière inéquation devient, puisque f est une bifonction d'équilibre,

$$f(x^k, y^k) + \frac{1}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle] \leq 0.$$

Ainsi, utilisant les deux dernières inégalités,

$$f(x^k, y^k) + \tfrac{1}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle]$$

$$\leq f(x^k, y^k) + \tfrac{\alpha}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle]$$

Vu qu'en outre, par la convexité forte de G,

$$G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle \geq 0,$$

nous déduisons que soit $G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle = 0$, soit $\alpha \geq 1$. Le premier cas implique que $x^k = y^k$, puisque G est fortement convexe. D'où, puisque par hypothèse $\alpha \in]0,1[$, les deux cas mènent à une contradiction.

2. Par la première assertion, il existe un entier fini strictement positif m tel que l'inégalité du Pas 2.1 tient. Puisque $f(x,.)$ est convexe, nous avons, étant donné que f est une bifonction d'équilibre,

$$0 = f(z^k, z^k) = f(z^k, [1-\theta_k]x^k + \theta_k y^k) \leq (1-\theta_k) f(z^k, x^k) + \theta_k f(z^k, y^k)$$

ou, de manière équivalente,

$$f(z^k, x^k) \geq \theta_k [f(z^k, x^k) - f(z^k, y^k)].$$

Utilisant l'inégalité du Pas 2.1, et vu que G est fortement convexe et que $y^k \neq x^k$, la dernière inéquation donne que

$$f(z^k, x^k) \geq \frac{\alpha\theta_k}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle] > 0.$$

3. Supposons par l'absurde que $0 \in \partial_2 f(z^k, x^k)$. Etant donné que $f(z^k,.)$ est convexe et que $f(x,.)$ admet un sous-différentiel, l'inclusion $0 \in \partial_2 f(z^k, x^k)$ implique que

$$f(z^k, x^k) \leq f(z^k, y) \ \forall y \in K.$$

Substituant $y = z^k \in K$ dans la dernière inéquation, nous obtenons $f(z^k, x^k) \leq 0$, ce qui contredit la précédente assertion.

Lemme 3.4.5 *Si $f(x,.)$ est convexe et admet un sous-différentiel, alors les assertions suivantes tiennent.*

1. *Pour toute solution x^* du (DEP), on a*

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2 - \gamma_k(2-\gamma_k)(\sigma_k \|g^k\|)^2.$$

2. $\sum_{k=0}^{\infty} \gamma_k(2-\gamma_k)(\sigma_k\|g^k\|)^2 < \infty.$
3. *Supposons que l'algorithme ne se termine pas. Alors, si ,en outre, $f(x,.)$ est finie sur un ensemble ouvert Ω contenant K et $f(\cdot, y)$ est continue, la suite $\{g^k\}$ est bornée.*

Preuve.

1. Prenons n'importe quel $x^* \in K^d$. Par la définition de x^{k+1} au Pas 3 qui est telle qu'en posant $w^k := x^k - \gamma_k \sigma_k g^k$, $x^{k+1} =: P_k(w^k)$, et par la propriété de quasi non expansivité par rapport à K de P_k, nous avons que

$$\|x^{k+1} - x^*\|^2$$

$$= \|P_k(w^k) - x^*\|^2$$

$$\leq \|w^k - x^*\|^2$$

$$= \|x^k - \gamma_k \sigma_k g^k - x^*\|^2$$

$$= \|x^k - x^*\|^2 - 2\gamma_k \sigma_k \langle g^k, x^k - x^* \rangle + (\gamma_k \sigma_k \|g^k\|)^2.$$

Etant donné que $f(z^k, .)$ est convexe et que $f(x, .)$ admet un sous-différentiel, l'inclusion $g^k \in \partial_2 f(z^k, x^k)$ implique

$$\langle g^k, x^k - x^* \rangle \geq f(z^k, x^k) - f(z^k, x^*).$$

D'autre part, puisque $x^* \in K^d$, nous avons $f(z^k, x^*) \leq 0$. Donc, en utilisant en outre le Pas 3 et la seconde assertion du précédent lemme,

$$\langle g^k, x^k - x^* \rangle \geq f(z^k, x^k) = \sigma_k \|g^k\|^2 > 0.$$

En utilisant les relations démontrée plus haut,

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2 - \gamma_k(2 - \gamma_k)(\sigma_k \|g^k\|)^2 \; \forall x^* \in K^d.$$

2. En appliquant la dernière inéquation pour tout k de 0 à m, nous obtenons

$$\sum_{k=0}^{m} \gamma_k(2 - \gamma_k)(\sigma_k \|g^k\|)^2 \leq \|x^0 - x^*\|^2 - \|x^{m+1} - x^*\|^2.$$

Puisque la suite positive $\{\|x^m - x^*\|\}_{m \geq 0}$ est décroissante par la première assertion et la première hypothèse sur γ_k, et donc convergente, en passant à la limite $m \to \infty$, nous obtenons

$$\sum_{k=0}^{\infty} \gamma_k(2 - \gamma_k)(\sigma_k \|g^k\|)^2 < \infty$$

3. Observons tout d'abord que $\{y^k\}$ est bornée. En effet, puisque y^k est l'unique solution du programme fortement convexe du Pas 1 dont la fonction objectif est continue et l'ensemble réalisable est constant, par le théorème maximal (voir Proposition 23 dans [1]), l'opérateur $s(x^k) : x^k \hookrightarrow y^k$ est continu. Puisque $\{x^k\}$ est bornée par la première assertion, $\{y^k\}$ est bornée, et, par conséquent, $\{z^k\}$ est également bornée, puisque z^k est combinaison convexe de x^k et y^k. Donc, par passage au théorème de Bolzano-Weierstrass et donc à une sous suite de $\{z^k\}$ et à son point d'adhérence z^*, $\lim_{k\to\infty} z^k = z^*$ où la sous-suite est, sans perte de généralité, indexée comme la suite. Par continuité par rapport à la première variable de la fonction convexe $f(z^k, .)$, la suite $\{f(z^k, .)\}$ converge simplement vers $f(z^*, .)$. Etant donné qu'en outre, f est convexe et finie par rapport à la seconde variable, et que $g^k \in \partial_2 f(z^k, z^k)$ par le Pas 3, nous pouvons déduire que $\{g^k\}$ est bornée

Théorème 3.4.2 *En plus des hypothèses des deux derniers lemmes, nous supposons que $f(x, .)$ est continue. Alors*

1. *la suite $\{x^k\}$ est bornée et chacun de ses points d'adhérence est une solution du (PEP).*
2. *si $K^d = K^*$ (en particulier, si f est pseudo monotone), alors la suite $\{x^k\}$ toute entière converge vers une solution du (PEP). En outre, si $\gamma_k = \gamma \in]0, 2[$ pour tout $k \geq 0$, alors*

$$lim\ inf_{k\to\infty} \left(\sigma_k \|g^k\| \sqrt{k+1}\right) = 0.$$

Preuve. Preuve identique à celle du théorème de convergence du premier algorithme de recherche linéaire.

Remarque 3.4.3 *D'après la première assertion du premier lemme relatif à l'algorithme qui nous occupe, la procédure de recherche linéaire du Pas 2 est toujours finie. Cependant, nous pouvons éviter la procédure de recherche linéaire si f satisfait la condition de type Lipschitz avec les deux constantes $c_1 > 0$ et $c_2 > 0$, c'est-à-dire si*

$$f(x,y) + f(y,z) \geq f(x,z) - c_1\|y-x\|^2 - c_2\|z-y\|^2 \ \forall x,y,z \in K.$$

Dans ce cas, nous pouvons choisir les constantes ρ et θ_k (pour tout $k \geq 0$) telles que

$$0 < \rho < \frac{(1-\alpha)\beta}{2c_2} \quad et \quad 0 < \theta_k \leq \sqrt{\frac{(1-\alpha)\beta - 2\rho c_2}{2\rho c_1}}.$$

La règle du Pas 2.1 tient alors toujours.

En effet, utilisant la condition de type Lipschitz avec $x = z^k$, $y = x^k$ et $z = y^k$, on obtient que

$$f(z^k, x^k) - f(z^k, y^k) \geq -f(x^k, y^k) - c_1\|x^k - z^k\|^2 - c_2\|y^k - x^k\|^2.$$

Etant donné que y^k est une solution du problème fortement convexe du Pas 1 qui est inchangé à une constante près, et que f est une fonction d'équilibre de sorte que $f(x^k, x^k) = 0$, nous avons

$$f(x^k, y^k) + \tfrac{1}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle] \leq 0$$

$$\Leftrightarrow -f(x^k, y^k) \geq \tfrac{1}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle].$$

Utilisant les deux dernières relations et un artifice de calcul, on obtient

$$f(z^k, x^k) - f(z^k, y^k) - \tfrac{\alpha}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle]$$

$$\geq -c_1\|x^k - z^k\|^2 - c_2\|y^k - x^k\|^2 + \tfrac{(1-\alpha)}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle].$$

D'autre part, étant donné que $z^k = (1-\theta_k)x^k + \theta_k y^k$, ce qui implique $\|x^k - z^k\|^2 = \theta_k^2\|x^k - y^k\|^2$, et que G est fortement convexe de module β, il suit de la dernière inégalité que

$$f(z^k, x^k) - f(z^k, y^k) - \tfrac{\alpha}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle]$$

$\geq [\frac{(1-\alpha)\beta}{2\rho} - c_1\theta_k^2 - c_2]\|y^k - x^k\|^2.$

Après résolution du système quadratique $\frac{(1-\alpha)\beta}{2\rho} - c_1\theta_k^2 - c_2 \geq 0$, *nous pouvons affirmer qu'en choisissant* $0 < \rho < \frac{(1-\alpha)\beta}{2c_2}$ *et* $0 < \theta_k \leq \sqrt{\frac{(1-\alpha)\beta - 2\rho c_2}{2\rho c_1}} < 1$ *(pour tout* $k \geq 0$*), alors*

$$f(z^k, x^k) - f(z^k, y^k) \geq \frac{\alpha}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle],$$

ce qui prouve cette remarque.

3.5 Les inéquations variationnelles mixtes.

Dans cette section, nous discutons d'applications des algorithmes proposés à l'inéquation variationnelle suivante, notée (MVIP),

trouver $x^* \in K$, $v^* \in F(x^*)$ tels que

$$\langle v^*, x - x^* \rangle + \varphi(x) - \varphi(x^*) \geq 0 \text{ pour tout } x \in K$$

où $F : \mathbb{R}^n \to \mathbb{R}^n$ est une application multivoque, et $\varphi : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ est une fonction fermée et convexe propre. Nous supposons que $F(x)$ est non vide et compacte $\forall x$, et que $K \subseteq dom(\varphi)$, où $dom(\varphi)$ représente le domaine effectif de φ.

Pour toute paire $x, y \in K$, en posant

$$f(x,y) := max_{u \in F(x)} \langle u, y - x \rangle + \varphi(y) - \varphi(x),$$

nous pouvons vérifier aisément que x^* est une solution du (MVIP) si et seulement s'il est une solution du (PEP). En effet, d'une part,

x^* solution du (MVIP)

$\Leftrightarrow \exists x^* \in K, \exists v^* \in F(x^*)$ tels que $\langle v^*, x - x^* \rangle + \varphi(x) - \varphi(x^*) \geq 0$ pour tout $x \in K$

$\Rightarrow \exists x^* \in K$ tel que $max_{u \in F(x^*)}[\langle u, x - x^* \rangle] + \varphi(x) - \varphi(x^*) \geq 0$ pour tout $x \in K$

$\Leftrightarrow \exists x^* \in K$ tel que $f(x^*, x) \geq 0$ pour tout $x \in K$

$\Leftrightarrow x^*$ solution du (PEP) ;

et, d'autre part,

x^* solution du (PEP)

$\Leftrightarrow \exists x^* \in K$ tel que $f(x^*, x) \geq 0$ pour tout $x \in K$

$\Leftrightarrow \exists x^* \in K$ tel que $max_{u \in F(x^*)}[\langle u, x - x^* \rangle] + \varphi(x) - \varphi(x^*) \geq 0$ pour tout $x \in K$

$\Rightarrow \exists x^* \in K, \exists v^* \in F(x^*)$ tels que $\langle v^*, x - x^* \rangle + \varphi(x) - \varphi(x^*) \geq 0$ pour tout $x \in K$

$\Leftrightarrow x^*$ solution du (MVIP).

Nous avons besoin des définitions qui suivent.

Définition 3.5.1 *L'opérateur F est dit*

1. *φ-pseudo monotone si et seulement si $\forall x, y \in K$ et $\forall u \in F(x), v \in F(y)$, nous avons*

$$\langle u, y - x \rangle + \varphi(y) - \varphi(x) \geq 0 \text{ implique } \langle v, y - x \rangle + \varphi(y) - \varphi(x) \geq 0;$$

2. *Lipschitz continu de module L si et seulement si $\forall x, y \in K$, nous avons*

$$sup_{u \in F(x)}\ inf_{v \in F(y)}\ \|u - v\| \leq L\|x - y\|.$$

Si $h(A, B)$ représente la distance de Hausdorff entre les deux ensembles A et B, c'est-à-dire si $h(A, B) = max\{sup_{x \in A}\ d(x, B), sup_{x \in B}\ d(x, A)\}$ où $d(x, A) = inf_{a \in A}\ \|x - a\|$, alors cette définition signifie que, $\forall x, y \in K$, nous avons

$$h[F(x), F(y)] \leq L\|x - y\|.$$

Comme nous l'avons vu dans l'exemple de la section précédente, la condition de type Lipschitz n'implique pas la continuité de f. Réciproquement, cependant, si $f(x, y) := max_{u \in F(x)}\langle u, y - x \rangle + \varphi(y) - \varphi(x)$ et si F est Lipschitz continu de module L, alors f satisfait la condition de type Lipschitz, comme l'établit le prochain lemme. C'est donc la raison pour laquelle une telle condition est appelée de type Lipschitz.

Lemme 3.5.1 *Soit f définie comme plus haut. Les assertions qui suivent tiennent.*

1. *Si F et φ sont continus, et si $F(x)$ est compact pour tout $x \in K$, alors f est continue.*
2. *Si F est φ-pseudo monotone, alors f est pseudo monotone.*
3. *Si F est Lipschitz continu de module L, alors, pour tout $\mu > 0$,*

$$f(x, y) + f(y, z) \geq f(x, z) - \frac{L\mu}{2}\|x - y\|^2 - \frac{L}{2\mu}\|y - z\|^2\ \forall x, y, z \in K.$$

Preuve.

1. Suit du théorème maximal (voir Proposition 23 dans [1]).
2. Suit immédiatement des définitions de f, de φ-pseudo monotonicité et de monotonicité en notant qu'avec $x, y \in K$,

$$\max_{u \in F(x)}\ [\langle u, y - x \rangle + \varphi(y) - \varphi(x)] \geq 0$$

$\Leftrightarrow \exists u' \in F(x)$ tel que $\langle u', y - x\rangle + \varphi(y) - \varphi(x) \geq 0$

$\Rightarrow \exists v' \in F(y)$ tel que $\langle v', y - x\rangle + \varphi(y) - \varphi(x) \geq 0$ par hypothèse

$\Leftrightarrow \max_{v\in F(y)} \left[\langle v, y - x\rangle + \varphi(y) - \varphi(x)\right] \geq 0$

3. Supposons que F est Lipschitz continu de module L. Soient $x, y, z \in K$. Pour tous $u \in F(x)$ et $\epsilon > 0$, puisque F est Lipschitz continu de module L, par définition, il existe $v \in F(y)$ tel que $\|u-v\| \leq L\|x-y\|+\epsilon$. Donc

   $\langle u, z - x\rangle - \max_{w'\in F(y)}\langle w', z - y\rangle - \max_{w\in F(x)}\langle w, y - x\rangle$

   $\leq \langle u, z - x\rangle - \langle v, z - y\rangle - \langle u, y - x\rangle$

   $= \langle u, z - y\rangle - \langle v, z - y\rangle$

   $= \langle u - v, z - y\rangle$

   $\leq \|u - v\|\|z - y\|$ par l'inégalité de Cauchy-Schwarz

   $\leq (L\|x - y\| + \epsilon)\|z - y\|.$

   Puisque $\epsilon > 0$ et $u \in F(x)$ sont arbitraires, en passant au $\max_{u\in F(x)}$ et à la limite $\epsilon \to 0$ de part et d'autre de l'inégalité qui précède, nous déduisons que

   $$f(x, z) - f(y, z) - f(x, y) \leq L\|x - y\|\|z - y\|$$

   $$\Leftrightarrow f(x, y) + f(y, z) \geq f(x, z) + (-\frac{1}{2}L)2\|x - y\|\|y - z\|.$$

   Utilisant l'inégalité bien connue $\frac{a^2}{\mu} + \mu b^2 \geq 2ab$, vraie pour tout $a, b \in$ IR et $\mu > 0$, résultant de la positivité de $\frac{1}{\mu}(b\mu - a)^2$, nous obtenons la thèse.

Notons que quand F est Lipschitz continu et univalué, l'algorithme de l'extragradient vu plus tôt dans ce chapitre, avec $f(x, y) := \langle F(x), y - x\rangle$, $\varphi \equiv 0$, devient l'algorithme bien connu de l'extragradient pour des inéquations variationnelles (voir chapitre précédent). Quand F est continu, les algorithmes de recherche linéaire s'accordent avec l'algorithme de projection sur un hyperplan (voir, de nouveau, chapitre précédent). Lorsque F est multivalué, par le lemme qui précède, en théorie, chacun des algorithmes présentés plus tôt peut être appliqué. En pratique, lorsque l'on implémente les algorithmes, il est difficile de choisir une approximation de $f(x, y)$.

3.6 Un problème particulier d'équilibre.

Dans cette section, nous illustrons les algorithmes proposés au moyen d'une classe du problème d'équilibre (PEP), pour laquelle K est un ensemble polyhèdral convexe donné par

$$K := \{x \in \mathbb{R}^n \mid Ax \leq b\},$$

et pour laquelle la bifonction d'équilibre $f : K \times K \to \mathbb{R} \cup \{\infty\}$ est de la forme

$$f(x, y) := \langle F(x) + Qy + q, y - x \rangle,$$

avec $F : K \to \mathbb{R}^n$, $Q \in \mathbb{R}^{n \times n}$ une matrice symétrique semi-définie positive, et $q \in \mathbb{R}^n$. Etant donné que Q est symétrique semi-définie positive, et donc que $\nabla_2^2 f(x, .)$ est semi-définie positive pour tout $x \in K$ fixé, $f(x, .)$ est convexe pour tout $x \in K$ fixé.

Pour cette classe, nous avons les résultats qui suivent.

Lemme 3.6.1 *Si $F : K \to \mathbb{R}^n$ est fortement monotone de module τ, alors*

1. *f est monotone quand $\tau \geq \|Q\|$.*
2. *f est fortement monotone de module $\tau - \|Q\|$ quand $\tau > \|Q\|$.*

Preuve.

1. Par définition de f, nous avons

   $$\begin{aligned}
   &f(x, y) + f(y, x) \\
   &= \langle F(x) + Qy + q, y - x \rangle - \langle F(y) + Qx + q, y - x \rangle \\
   &= \langle Q(y - x), y - x \rangle - \langle F(y) - F(x), y - x \rangle.
   \end{aligned}$$

   Notons que

   $\langle Q(y-x), y-x \rangle \leq \|Q(y-x)\| \|y-x\|$ par l'inégalité de Cauchy-Schwarz

   $\leq \|Q\| \|y - x\|^2$ par les propriétés des normes matricielles.

   Etant donné que F est fortement monotone de module τ, c'est-à-dire que

   $$-\langle F(y) - F(x), y - x \rangle \leq -\tau \|y - x\|^2,$$

   nous avons, en utilisant les relations qui précèdent, que

   $$f(x, y) + f(y, x) \leq 0 \text{ quand } \tau \geq \|Q\|.$$

2. De même, nous avons que

   $f(x, y) + f(y, x) \leq -(\tau - \|Q\|) \|y - x\|^2$ avec $\tau - \|Q\| > 0$ quand $\tau > \|Q\|$.

Lemme 3.6.2 *Si F est Lipschitz continu de module L, c'est-à-dire si*

$$\|F(y) - F(x)\| \leq L\|y - x\| \ \forall x, y \in K,$$

alors f satisfait la constante de type Lipschitz, c'est-à dire

$$f(x,y) + f(y,z) \geq f(x,z) - c_1\|y - x\|^2 - c_2\|z - y\|^2 \ \forall x, y, z \in K,$$

pour tous $c_1, c_2 > 0$ vérifiant

$$2\sqrt{c_1 c_2} \geq L + \|Q\|.$$

Preuve. Pour tous $x, y, z \in K$, nous avons

$$f(x,y) + f(y,z) - f(x,z)$$

$$= \langle F(x) + Qy + q, y - x\rangle + \langle F(y) + Qz + q, z - y\rangle - \langle F(x) + Qz + q, z - x\rangle$$

$$= \langle F(y) - F(x), z - y\rangle + \langle Q(y - z), y - x\rangle.$$

Appliquant l'inégalité de Cauchy-Schwarz, nous avons

$$\langle Q(y - z), y - x\rangle$$

$$= -\langle Q(z - y), y - x\rangle$$

$$\geq -\|Q(z - y)\|\|y - x\|$$

$$\geq -\|Q\|\|z - y\|\|y - x\| \text{ par les propriétés des normes matricielles,}$$

et

$$\langle F(y) - F(x), z - y\rangle$$

$$= -\langle F(y) - F(x), y - z\rangle$$

$$\geq -\|F(y) - F(x)\|\|y - z\|.$$

Puisque F est Lipschitz continu de module L, nous pouvons écrire

$$-\|F(y) - F(x)\|\|z - y\| \geq -L\|y - x\|\|z - y\|$$

Donc, des trois inégalités démontrées ci-avant, il suit que

$$f(x,y) + f(y,z) - f(x,z) \geq -(L + \|Q\|)\|y - x\|\|z - y\|.$$

Dès lors, étant donné que, par hypothèse, $-(L + \|Q\|) \geq -2\sqrt{c_1 c_2}$, et que, par positivité de $(\sqrt{c_1}\|y - x\| - \sqrt{c_2}\|z - y\|)^2$, $-2\sqrt{c_1 c_2}\|y - x\|\|z - y\| \geq -c_1\|y - x\|^2 - c_2\|z - y\|^2$, nous avons

$$-(L + \|Q\|)\|y - x\|\|z - y\| \geq -c_1\|y - x\|^2 - c_2\|z - y\|^2.$$

Donc

$$f(x,y) + f(y,z) \geq f(x,z) - c_1\|y-x\|^2 - c_2\|z-y\|^2 \ \forall x,y,z \in K.$$

Dans le cas spécial où F est un opérateur linéaire de la forme $F(x) = Px$ avec $P \in \mathbb{R}^{n\times n}$, la fonction f définie en début de section prend la forme

$$f(x,y) = \langle Px + Qy + q, y - x\rangle.$$

Propriété 3.6.1 *Supposons que les matrices P et Q sont choisies de sorte que Q soit symétrique semi-définie positive et que $Q - P$ soit semi-définie négative. Alors*

1. *f est monotone, $f(\cdot, y)$ est continue pour tout $y \in K$ fixé et $f(x, .)$ est différentiable et convexe pour tout $x \in K$ fixé.*
2. *Pour tous $x, y, z \in K$, on a*

$$f(x,y) + f(y,z) \geq f(x,z) - c_1\|z-y\|^2 - c_2\|y-x\|^2$$

où $c_1 = c_2 = \frac{1}{2}\|P - Q\|$.

Preuve.

1. Pour tous $x, y \in K$, puisque $Q - P$ est semi-définie négative, nous avons, par définition de f,

$$f(x,y) + f(y,x)$$
$$= \langle Px + Qy + q, y - x\rangle - \langle Py + Qx + q, y - x\rangle$$
$$= \langle (Q-P)(y-x), y - x\rangle$$
$$\leq 0.$$

D'où, f est monotone. Clairement, $f(\cdot, y)$ est continue, $f(x, .)$ est différentiable, et étant donné que Q est symétrique semi-définie positive, et donc que $\nabla_2^2 f(x, .)$ est semi-définie positive pour tout $x \in K$ fixé, $f(x, .)$ est convexe pour tout $x \in K$ fixé.

2. Observons que, pour tous $x, y, z \in K$,

$$f(x,y) + f(y,z) - f(x,z)$$
$$= \langle Px + Qy + q, y - x\rangle + \langle Py + Qz + q, z - y\rangle - \langle Px + Qz + q, z - x\rangle$$
$$= \langle Px + Qy, y - x\rangle + \langle Py + Qz, z - y\rangle - \langle Px + Qz, z - x\rangle + \langle q, y - x + z - y - (z - x)\rangle$$
$$= \langle Px + Qy, y - x\rangle + \langle Py + Qz, z - y\rangle - \langle Px + Qz, z - x\rangle$$
$$= \langle Px, y - x - (z-x)\rangle + \langle Qz, z - y - (z - x)\rangle + \langle Qy, y - x\rangle + \langle Py, z - y\rangle$$

$= \langle Px, y - z\rangle + \langle Qz, x - y\rangle + \langle Qy, y - x\rangle + \langle Py, z - y\rangle$

$= \langle P(y - x), z - y\rangle + \langle Q(z - y), x - y\rangle$ par linéarité de P et Q

$= \langle P(y - x), z - y\rangle + \langle Q^T(x - y), z - y\rangle$

$= \langle P(y-x), z-y\rangle + \langle Q(x-y), z-y\rangle$ puisque $Q = Q^T$ par symétrie de Q

$= \langle (P - Q)(y - x), z - y\rangle$.

Il s'ensuit alors que $f(x, y) + f(y, z) - f(x, z)$

$= \langle (P - Q)(y - x), z - y\rangle$

$= -\langle (P - Q)(x - y), z - y\rangle$

$\geq -\|(P - Q)(x - y)\|\|z - y\|$ par l'inégalité de Cauchy-Schwarz

$\geq -\|P - Q\|\|x - y\|\|z - y\|$ par les propriétés des normes matricielles

$= \frac{\|P-Q\|}{2}[-2\|x - y\|\|z - y\|]$ en utilisant un artifice de calcul

$\geq \frac{\|P-Q\|}{2}[-\|x - y\|^2 - \|z - y\|^2]$ par positivivé de $(\|y - x\| - \|z - y\|)^2$.

En posant, par exemple, $c_1 = c_2 = \frac{1}{2}\|P - Q\|$, on obtient la thèse.

On peut utiliser l'algorithme de l'extragradient avec la fonction quadratique de régularisation $G(x) := \frac{1}{2}\|x\|^2$ pour résoudre le (PEP) avec $f(x, y) := \langle Px + Qy + q, y - x\rangle$ et $K := \{x \in \mathbb{R}^n \mid Ax \leq b\}$. Dans ce cas, le programme fortement convexe du Pas 1, qui est inchangé à une constante près, est de la forme

$$\min_{y\in K} \ [\rho f(x, y) + \frac{1}{2}\|y\|^2 - \langle x, y - x\rangle]$$

$$\Leftrightarrow \min_{y\in K} \ [\rho f(x, y) + \frac{1}{2}\|y\|^2 - \langle x, y\rangle + \frac{1}{2}\|x\|^2 + \frac{1}{2}\|x\|^2]$$

$$\Leftrightarrow \min_{y\in K} \ [\rho f(x, y) + \frac{1}{2}\|y - x\|^2].$$

On peut également utiliser le premier algorithme de recherche linéaire avec les mêmes bifonction d'équilibre et fonction quadratique de régularisation qu'avant. Dans cet algorithme, au Pas 3 de la kième itération, on choisira x^{k+1} comme étant la projection euclidienne de $x^k - \gamma_k\sigma_k g^k$ sur K. Il est bien connu que calculer une telle projection mène à un programme quadratique convexe.

Dans ce cas particulier, le second algorithme de recherche linéaire se ramène à celui qui suit (voir prochain lemme).

Algorithme 3.6.1

Données : $x^0 \in K$, $\epsilon > 0$, $\alpha \in]0,1[$, $\theta \in]0,1[$ *et* $\rho > 0$.

Pas 0 : *Poser* $k := 0$.

Pas 1 : *Résoudre le programme quadratique fortement convexe*

$$min_{y\in K}\ [\frac{1}{2}\langle Hy, y\rangle + \langle h(x^k), y\rangle],$$

où $H = 2\rho Q + I$ *et* $h(x^k) = [\rho(P-Q) - I]x^k + \rho q$, *pour obtenir son unique solution* y^k.

Si $\|y^k - x^k\| \leq \epsilon$, *alors stop :* x^k *est une solution approximative du (PEP). Sinon, aller au Pas 2.*

Pas 2 : *Choisir* $\theta_k \in]0,1[$ *tel que*

$$0 < \theta_k \leq min\ \{\frac{u(x^k, y^k)}{v(x^k, y^k)}, \theta\}$$

où

$$u(x^k, y^k) = \frac{1}{2\rho}\langle(2\rho P - \alpha I)x^k + 2\rho Q - \alpha I)y^k + 2\rho q, x^k - y^k\rangle$$

et

$$v(x^k, y^k) = \langle(P-Q)(x^k - y^k), x^k - y^k\rangle.$$

Poser $z^k := (1-\theta_k)x^k + \theta_k y^k$, *et aller au Pas 3.*

Pas 3 : *Prendre* $g^k := (P-Q)z^k + 2Qx^k + q$ *et*

$$\sigma_k = \frac{\langle Pz^k + Qx^k + q, x^k - z^k\rangle}{\|g^k\|^2}.$$

Résoudre le programme quadratique convexe

$$min_{y\in K}\ [\frac{1}{2}\langle y, y\rangle + \langle c(x^k, z^k), y\rangle],$$

où $c(x^k, z^k) = \gamma_k \sigma_k g^k - x^k$, *pour obtenir son unique solution* x^{k+1}.

Pas 4 : *Poser* $k := k+1$ *; revenir au Pas 1.*

Lemme 3.6.3 *Soit* $f(x,y) := \langle Px + Qy + q, y - x\rangle$. *Supposons que la matrice* Q *est symétrique semi-définie positive et que la matrice* $P - Q$ *est définie positive. Alors, le second algorithme de recherche linéaire se ramène à celui-ci.*

Preuve. Tout d'abord, montrons que le programme fortement convexe du second algorithme de recherche linéaire qui est inchangé à une constante près, se ramène au premier programme quadratique fortement convexe de cet algorithme. En effet, nous avons,

$$\min_{y\in K}\ \left(f(x^k, y) + \frac{1}{\rho}[\frac{1}{2}\|y\|^2 - \langle x^k, y - x^k\rangle]\right)$$

$$\Leftrightarrow \min_{y\in K} \left(f(x^k, y) + \frac{1}{\rho}[\frac{1}{2}\|y\|^2 - \langle x^k, y\rangle + \frac{1}{2}\|x^k\|^2 + \frac{1}{2}\|x^k\|^2] \right)$$

$$\Leftrightarrow \min_{y\in K} \, [f(x^k, y) + \frac{1}{2\rho}\|y - x^k\|^2]$$

où

$f(x^k, y) + \frac{1}{2\rho}\|y - x^k\|^2$

$= \langle Px + Qy + q, y - x\rangle + \frac{1}{2\rho}\|y - x^k\|^2$ par définition de f

$= \frac{1}{2\rho}\langle (2\rho Q + I)y, y\rangle + \frac{1}{\rho}\langle [\rho(P - Q) - I]x^k + \rho q, y\rangle + \frac{1}{2\rho}\|x^k\|^2 - \langle Px^k + q, x^k\rangle$ par réarrangement des termes (en utilisant la symétrie de Q).

Donc, si nous prenons $H = 2\rho Q + I$ et $h(x^k) = [\rho(P - Q) - I]x^k + \rho q$, alors il est aisé de voir que le programme fortement convexe du second algorithme de recherche linéaire se ramène au premier programme quadratique fortement convexe de cet algorithme, inchangé à une constante et une constante multiplicative près.

Ensuite, observons que les procédures de recherche linéaire du Pas 2 du second algorithme de recherche linéaire et de cet algorithme sont équivalentes. En effet, nous avons

$f(z^{k,m}, x^k) - f(z^{k,m}, y^k) - \frac{\alpha}{\rho}[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle]$

$= \langle Pz^{k,m} + Qx^k + q, x^k - z^{k,m}\rangle - \langle Pz^{k,m} + Qy^k + q, y^k - z^{k,m}\rangle - \frac{\alpha}{\rho}[\frac{1}{2}\|y^k\|^2 - \frac{1}{2}\|x^k\|^2 - \langle x^k, y^k - x^k\rangle]$ par définition de f

$= \langle Pz^{k,m} + Qx^k + q, x^k - z^{k,m}\rangle - \langle Pz^{k,m} + Qy^k + q, y^k - z^{k,m}\rangle - \frac{\alpha}{\rho}[\frac{1}{2}\|y^k\|^2 - \frac{1}{2}\|x^k\|^2 - \langle x^k, y^k\rangle + \|x^k\|^2]$

$= \langle Pz^{k,m} + Qx^k + q, x^k - z^{k,m}\rangle - \langle Pz^{k,m} + Qy^k + q, y^k - z^{k,m}\rangle - \frac{\alpha}{\rho}[\frac{1}{2}\|y^k\|^2 - \langle x^k, y^k\rangle + \frac{1}{2}\|x^k\|^2]$

$= \langle Pz^{k,m} + Qx^k + q, x^k - z^{k,m}\rangle - \langle Pz^{k,m} + Qy^k + q, y^k - z^{k,m}\rangle - \frac{\alpha}{2\rho}\|x^k - y^k\|^2$

$= \langle Px^k + Qy^k + q, x^k - y^k\rangle - \theta_k\langle (P - Q)(x^k - y^k), x^k - y^k\rangle - \frac{\alpha}{2\rho}\|x^k - y^k\|^2$ par définition de $z^{k,m}$ et réarrangement des termes

$= \frac{1}{2\rho}\langle 2\rho(Px^k + Qy^k + q) - \alpha(x^k - y^k), x^k - y^k\rangle - \theta_k\langle (P - Q)(x^k - y^k), x^k - y^k\rangle$
$\geq 0.$

En posant

$$u(x^k, y^k) = \frac{1}{2\rho}\langle (2\rho P - \alpha I)x^k + (2\rho Q + \alpha I)y^k + 2\rho q, x^k - y^k\rangle$$

et

$$v(x^k, y^k) = \langle (P - Q)(x^k - y^k), x^k - y^k \rangle,$$

il suit de la précédente inégalité que

$$u(x^k, y^k) \geq \theta_k v(x^k, y^k).$$

Etant donné que $x^k \neq y^k$ et que $P - Q$ est une matrice définie positive, $v(x^k, y^k) > 0$. Donc, nous pouvons choisir

$$0 < \theta_k \leq min \ \{\frac{u(x^k, y^k)}{v(x^k, y^k)}, \theta\} < 1,$$

ce qui prouve que les procédures de recherche linéaire du Pas 2 du second algorithme de recherche linéaire et de cet algorithme sont équivalentes.

Finalement, puisque $f(x, y) := \langle Px + Qy + q, y - x \rangle$ est différentiable, nous avons $g^k = \nabla_2 f(z^k, x^k) = (P - Q)z^k + 2Qx^k + q$ et $f(z^k, x^k) = \langle Pz^k + Qx^k + q, x^k - z^k \rangle$, ce qui implique que

$$\sigma_k := \frac{\langle Pz^k + Qx^k + q, x^k - z^k \rangle}{\|g^k\|^2}.$$

Utilisant le Pas 3 du second algorithme de recherche linéaire, nous obtenons

$$x^{k+1} = P_k(x^k - \gamma_k \sigma_k g^k),$$

ce qui montre que x^{k+1} est, par définition de la projection orthogonale, l'unique solution du problème

$$\min_{y \in K} \ [\|y - (x^k - \gamma_k \sigma_k g^k)\|^2]$$

$$\Leftrightarrow \min_{y \in K} \ [\|y\|^2 - 2\langle x^k, y \rangle + 2\gamma_k \sigma_k \langle g^k, y \rangle + \|x^k - \gamma_k \sigma_k g^k\|^2],$$

qui peut se réécrire, puisqu'il est inchangé à une constante et une constante multiplicative près,

$$\min_{y \in K} \ [\frac{1}{2}\|y\|^2 + \langle c(x^k, y^k), y \rangle]$$

où $c(x^k, y^k) := \gamma_k \sigma_k g^k - x^k$. Ceci achève donc la preuve.

Bibliographie

[1] J.P. Aubin and I. Ekeland, Applied Nonlinear Analysis, Wiley, New-York, 1984.

[2] E. Blum and W. Oettli, *From optimization and variational inequality to equilibrium problems*, The Mathematics Student (Volume 63), 1994, pp. 127-149.

[3] Dinh Quoc Tran, Muu Le Dung and V.H. Nguyen, *Extragradient algorithms extended to equilibrium problems*, Journal of Optimization (à paraître).

[4] F. Facchinei and J.S. Pang, *Finite-Dimensional Variational Inequalities and Complementary Problems* (Chapter 12), Springer-Verslag, New-York, 2003.

[5] G. Mastroeni, *On auxiliary principle for equilibrium problems*, Publicatione del Dipartimento di Mathematica dell'Universita di Pisa (Volume 3), 2000, pp. 1244-1258.

[6] V.H. Nguyen, *An introduction to variational inequalities and related problems*, FUNDP Namur, 2003.

[7] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.